Modelling the effect of HIV on age-specific incidence of active TB disease: a comparison between Taiwan and Cape Town Metropole

by

Dietrich Maximilian Albert **WINKLER**

Thesis presented in partial fulfilment of the
academic requirements for the degree of
Master of Science
at the Stellenbosch University

Supervisor : Doctor Farai **NYABADZA**
Department of Mathematical Sciences
University of Stellenbosch

September 2011

Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and has not previously, in its entirety or in part, been submitted at any university for a degree.

1 September 2011

Dietrich Maximilian Albert Winkler

Date

Dedication

I dedicate this project to my most wonderful parents, Mr Armin Dietrich Winkler, Mrs Jeanette Maria Winkler and also to my two sisters Anja and Siegrid.

Abstract

The dynamics of tuberculosis (TB) are largely dependent on age. In South Africa the dynamics of TB have also been influenced by HIV/AIDS. Studies have shown that the presence of active TB disease has increased 2 to 3 fold since the advent of HIV in sub-Saharan Africa. Determination of the effects of HIV on TB dynamics has been modelled through coinfection models that are not age-dependent. In this thesis we consider a TB model that is age-dependent. The model is fitted to age-specific data from Taiwan where HIV prevalence is assumed to be negligibly small. The Taiwan data is thus used as the baseline in our study. The same model is fitted to the Cape Town Metropole data. It is known that the prevalence of HIV in Cape Town Metropole is 15%. Its effect on TB is determined by comparing the parameter values from the two fits. The parameters are set as functions of age. The effective contact rate, a function of both age and time, was changed to fit our model to the notification rates of active TB disease cases for both regions. In Cape Town Metropole the incidence of active TB disease for children and the highly sexually active is considerably higher than for the correspondig ages of Taiwan. The incidence of active TB disease for the elderly in Taiwan is comparable with that of Cape Town Metropole. The results also reflect that higher incidence of active TB disease goes along with higher effective contact rates. Our results show that the age groups who have the highest incidence rates of active TB disease should have the highest contribution in the transmission of TB. Depending on the age group and the time considered, the effect of HIV on the prevalence of active TB disease was determined to be between 6 and 337 times higher than when it is absent. It is assumed that the difference in living conditions is negligibly small on the effect of TB.

Opsomming

Die dinamika van tuberkulose (TB) is hoofsaaklik afhanklik van ouderdom. In Suid-Afrika het MIV/VIGS ook die dinamika van TB beïnvloed. Studies in sub-Sahara Afrika, het getoon dat die aanwesigheid van aktiewe TB siekte, het sedert die koms van MIV met 2 tot 3 keer toegeneem. Bepaling van die uitwerkinge van MIV op TB dinamika is gemodelleer deur ko-infeksie modelle wat nie ouderdomsafhanklik is nie. In hierdie tesis beskou ons 'n TB model wat ouderdomsafhanklik is. Die model word gepas aan ouderdomspesifieke data van Taiwan waar aangeneem is dat MIV voorkoms ondeduidend is. Die Taiwan data word dus gebruik as die grondlyn van ons studie. Dieselfde model word gepas aan die Kaapstap Metropool data. Dit is bekend dat die voorkomsyfer van MIV in Kaapstap Metropool 15% is. MIV se uitwerking op TB word bepaal deur die parameterwaardes van beide passings met mekaar te vergelyk. Die parameters word vasgestel as funksies van ouderdom. Die effektiewe kontak koers, 'n funksie van beide ouderdom en tyd, word verander om ons model aan die kennisgewingskoerse gevalle van aktiewe TB siekte vir beide streke te pas. In die Kaapstad Metropool is die insidensie van aktiewe TB siekte vir kinders en die hoogs seksuele aktiewes aansienlik hoër as dié van die ooreenstemmende ouderdomme van Taiwan. Die insidensie van aktiewe TB siekte van bejaardes in Taiwan is vergelykbaar met dié van Kaapstad Metropool. Die resultate weerspieël dat hoër insidensie van aktiewe TB siekte met hoër effektiewe kontak koerse gepaard gaan. Ons resultate toon dat die ouderdomsgroepe wat die hoogste insidensie koerse van aktiewe TB siekte het, behoort die hoogste bydrag tot die oordrag van TB te hê. Afhangende van die ouderdomsgroep en die tyd onder beskouing, is bepaal dat die uitwerking van MIV op die voorkoms van aktiewe TB siekte tussen 6 en 337 keer hoër is, as wanneer dit afwesig is. Dit word aangeneem dat die verskil in lewensomstandighede weglaatbaar klein is op die uitwerking van TB.

Acknowledgements

First of all I would like to thank the Heavenly Father for giving me this wonderful opportunity to study at this level. I have much gratitude for the work my supervisor Dr Farai Nyabadza did. He provided me with very optimistic support and he gave the perfect guidance at appropriate times. I thank him for his big sacrifice supervising me as one of his students. I acknowledge the financial assistance of the National Research Foundation (NRF) and the South African Centre for Epidemiological Modelling and Analysis (SACEMA). Without their assistance all of this work would not have been possible.

Many thanks go to Dr Judy Caldwell and Dr Karen Jennings from City of Cape Town Health Department for providing me with appropriate data for my dissertation. Without their assistance my dissertation would have had less worth.

I have many thanks for my father Armin Winkler who sponsered my undergraduate studies which in turn made it possible for me to do this project at postgraduate level. I thank my mother Jeanette Winkler together with my two sisters Anja and Siegrid for their emotional support. Last but not least I would like to thank my collegues at the South African Centre for Epidemiological Modelling and Analysis (SACEMA) for their academic and emotional support regarding my dissertation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Mycobacterium tuberculosis (Mtb) is transmitted by airborne particles of bacilli. Tuberculosis (TB) infection results when the bacterium is deposited in the lungs of exposed persons [3, 19]. A small proportion (approximately 10%), see [17], of persons progress to active TB and become infectious within the first two years of infection, while the remainder remains latently infected. Latent infections are asymptomatic and do not contribute to TB transmission, however they may progress to active TB disease through either endogenous reactivation or exogenous re-infection.

Infection with Mtb thus leads to active TB disease by one of the three possibilities:

- primary progression after a recent infection,
- re-activation of a latent infection,
- exogenous re-infection of a previously infected individual.

The estimated effect of re-infection is can be determined by mathematical models by considering parameters that adhere to epidemiological data. Studies reflect that exogenous re-infection is important only when TB incidence is high. This is of particular importance in the context of South Africa. Whether one can develop disease fast or have latent infection depends on an individual's immune response to the infection and the virulence of

the pathogen when the person gets infected. Latently infected individuals can undergo reactivation. After a cure, a relapse can occur and an individual develops the active TB again [14]. Secondary infections by the same pathogenic agent that occur after primary infection can be defined as re-infection. Re-infection by pathogens is however independent of disease status (it can occur in both the latent and recovered states). There are thus two processes (reactivation and re-infection) that lead to the progression to disease for individuals who are latently infected. Reactivation is an immune selection driven process and re-infection is an interactive process with the pathogen (through diseased persons). Since the effect of these two processes is uncertain, DNA fingerprinting is used to explain the influence of these two mechanisms. Immune suppression usually caused by HIV infection leads to higher rates of exogenous re-infection [30]. Pulmonary TB, which refers to TB in the lungs is the most common form of TB [24]. Adults might suffer from this type of TB many years after a primary infection arising from either endogenous reactivation of the primary infection or a recent, exogenous re-infection [30].

Re-infection cannot occur among individuals with active TB disease or among those already at high risk of developing the first primary episode or exogenous disease (i.e. within 5 years of the initial infection) [45]. However, it is still unknown how treatment influences re-infection. Three basic scenarios arise with regard to TB treatment, leading to the following questions [19]: Firstly, does treatment of latent TB not affect the risk of developing active TB disease upon re-infection? Secondly, does treatment decrease susceptibility to re-infection? Thirdly, does treatment increase susceptibility to re-infection? If the diagnosis and treatment of TB is successful then the duration of active disease since the onset of the disease is more or less 2 years [2].

Three types of medical interventions are available to address TB [19]:

- vaccination to prevent infection,
- treatment to cure active TB disease,
- treatment of latent TB to prevent endogenous reactivation.

Bacille Calmette-Guerin (BCG) is the sole vaccine against TB and its efficacy varies from 0% to 80% depending on the population under consideration. The treatment to cure active TB disease involves short-course chemotherapy under proper management and also directly observed therapy which involves a regular drug supply together with a thorough monitoring system. HIV infected individuals who also have a latent TB infection have a high risk of developing active TB disease. Therefore, these individuals are treated with isoniazid for 6 to 12 months or with rifampicin and pyrazinamide for 2 months to reduce their chances of developing active TB [8].

1.2 Age-specific Incidence of TB

The scenario for Cape Town Metropole regarding the incidence of active TB disease holds as follows: within the 0-5 years age range the incidence of active TB disease is quite high, since young children are at higher risk of developing active disease once they are infected [44]. The infection is predominantly as a result of contacts with infectious adults within their households. Then within the 5-12 years age range, the incidence of active TB disease is surprisingly low. These are known as the "golden years". After 12 years the incidence of TB increases considerably as the age increases. Then from 35 years onwards the incidence of TB decreases as age increases [30]. In Taiwan on the other hand, the incidence of Taiwan was the highest for the 25-44 year age group (50.9%) during the 1957-1961 period, the highest for the 45-64 year age group (44.9%) during the 1977-1981 period and the highest for individuals aged 65 years or older (43.7%) during the 1997-2001 period [51]. However, in Taiwan the age distribution of active TB disease incidence for the 1957-1961 period is similar to that of Cape Town during the 2003-2009 period.

Re-infection and age need to be incorporated in TB models since they are two of the most important underlying factors that influence the dynamics of TB. When Vynnycky (1997) *et al.* [45] excluded active TB disease which arose from exogenous re-infection from their model, the fit to observed notification rates got worsened. This was because about 60% and 80% of the active disease incidence for the 40 and 60 year olds respectively during the 1950s in their study, was a result of exogenous re-infection.

1.3 Vaccination of TB

Bacillus of Calmette and Guérin (BCG) is currently the only vaccine in use against TB. Despite its wide usage, the effectiveness of BCG in preventing TB is unclear since it is estimated that BCG is between 0% and 80% effective [12, 19]. The protection that BCG vaccination provides against pulmonary TB, is therefore controversial [24]. BCG is a live, less intense strain of *Mycobacterium bovis* [6] and it is more involved in reducing the progression of disease rather than the risk of actual infection [18, 20]. In some populations the common use of BCG vaccine are closely related to the potential problem that vaccinated individuals will test positive for latent TB [12]. BCG may lead the immune system to respond stronger to the Tubercullin Skin Test (TST) after MTB infection since vaccination plays an important role in the routine immunological memory of MTB infection [6, 16]. Therefore, it becomes very difficult to obtain realistic estimates for the prevalence of disease in a population [12].

However, timing vaccination through BCG effectively, could have important implications for improving protection against TB disease [24]. Vaccinated individuals who get infected, have a higher probability of remaining latently infected over the long run and have a low risk of developing active TB disease [2]. Since the efficacy of the vaccine wanes over time, BCG vaccination served during infancy has little effect on the incidence of TB in individuals older than 15 years of age [46].

1.4 Problem Statement

Tuberculosis is a major problem in the Western Cape. The situation is exacerbated by HIV/AIDS, which has resulted in increased immuno compromisation. TB transmission is age-dependent and adults readily transmit TB to children. On the other hand children do not transmit TB [49]. There is a tendency for HIV infected children to develop TB. Before the start of the HIV epidemic in the Western Cape the TB notification rates have been the highest among the elderly. Currently, however the TB notification rates are the highest among the economically productive age groups. HIV negative individuals with active TB disease contribute more to transmission of TB than those who are HIV positive [26]. TB

data does not show the contribution of HIV across ages, yet the influence of HIV is very evident. Some models have included TB and HIV submodels coupled together [10, 29]. The parameters are usually regarded as constants. On the other hand, parameters linking the two can not be explicitly found or determined. Planning for interventions requires reasonable and realistic projections of the incidence across the ages as TB dynamics in adults is different from that in children with the former determining the latter. The parameter should thus be age-dependent as transmission differs across ages.

1.5 Approach

Our approach is to consider firstly a scenario in which TB dynamics is not driven by HIV. We therefore consider the case of Taiwan and fit our model to the Taiwan data. We determine the model parameters as we take them to be the model parameters for TB only dynamics. The same model is fitted to the Cape Town Metropole data. We then argue that the differences in the parameter values and data, show the potential contribution of HIV to TB dynamics without necessarily having to use co-infection models which are usually intractable. This is done using an age dependent TB model.

1.6 Motivation

The project is driven by the need to determine TB trends in Cape Town Metropole. We make projections of active TB disease incidences in the various age groups for the purpose of designing and planning of interventions. Moreover, we also need to quantify the contribution of HIV to active TB disease incidence in Cape Town Metropole. There is a general tendency for the age-specific incidence of tuberculosis to vary greatly within and between countries over time [35, 51]. This is because of problems in determining tuberculosis infection [45] and the effect of HIV which increases the annual incidence rates of active TB disease since 1990 by 2 to 3-fold in many countries of sub-Saharan Africa [26].

The emphasis will be on the transmission parameters between the considered age groups, hopefully to explain the clinical data on how the incidence is high during the infant stage with a decline in incidence of older children between 5-12 years of age, to an increase of

incidence for ages above 12 years. By means of mathematical modelling, a clearer picture can be obtained of the relative importance of the age-specific risks of developing active TB disease. Hence, a connection can be drawn between the current knowledge of the pathogenesis of tuberculosis and the past incidence of active TB disease data [45].

1.7 Aims and Objectives

The major aim of the project is to fit a simple age-structured TB model to age structured data from the Cape Town Metropole with the object of making projections that give the future trends of the TB epidemic in the Metropole. This is of particular importance in the designing and planning of interventions. This can be achieved through the following sub-aims:

- Fitting the model to data obtained from a population whose TB dynamics are not affected by HIV such as Taiwan.
- To obtain the model parameters specific to TB only.
- To fit the same model to data from the Cape Town Metropole.
- The model parameters are compared.
- The TB parameters that are most affected by HIV are determined and comparisons are made.
- To give pictorial presentations in the form of graphs.

1.8 Project's Description

The project will introduce continuous age structured models, which basically consists of an initial-boundary value problem of partial differential equations. The idea is to produce three-dimensional graphs with respect to active TB disease incidence, time and age, and make projections of active TB disease incidence. Firstly, the effective contact rates of different age groups that provide the best fit are a function of age and time. Therefore,

the model allows different effective contact rates between individuals of different ages. The other parameters of the model are only functions of age and their values are assumed and/or approximated from literature which was appropriate to more or less the settings of Cape Town Metropole and Taiwan, respectively. The effective contact rate, a function of both age and time was changed to fit our model to the notification rates of active TB disease cases for both settings. This results in smoother fits to the model.

1.8.1 Thesis Structure

The outline of the dissertation is as follows; Chapter 1 provides a general introduction regarding TB and related aspects involving age-structure. It also gives a brief comparison between the TB epidemic in Taiwan and Cape Town Metropole. In Chapter 2 we review some literature associated with TB in the Western Cape, TB models, age-structured models, fitting and some mathematical preliminaries. Chapter 3 presents a TB age-structured model. In Chapter 4 the model is fit to the Taiwan data while in Chapter 5 the model is fit to the Cape Town Metropole data. Chapter 6 provides the comparison of the Taiwan and Cape Town Metropole fit together with their associated parameters. Finally we discuss the implications.

Chapter 2

Literature Review

2.1 A Brief Introduction on TB models

Tuberculosis is a very complex disease and modellers approach TB modelling from different angles to address questions which biologists have related to TB dynamics. Apart from age, modellers also lay heavy emphasis on the relative importance of re-infection and reactivation regarding TB dynamics. The effect of reactivation increases in importance as age increases.

Feng *et al.* [17] stated that the longer an individual is latently infected with TB, the greater the immunity against the development of active TB disease, unless there are other diseases that attack the immune system. Therefore, age of infection and chronological age are important factors to take into consideration for disease progression. Uys *et al.* [42] and Aparicio *et al.* [3] agree with the importance of these two factors. The introduction of re-infection according to Feng *et al.* [17] increases the unpredictability of TB which makes it difficult to set up public health policies. The issues of socioeconomic conditions and overcrowding are discussed in relation to TB transmission. However, all these models did not take age-structure into account.

Gabriela *et al.* [18, 20] introduced the concept of a re-infection threshold as a benchmark to their model and investigated the effect of vaccination and transmission relative to the re-infection threshold. They explained how primary infection plays a role below the re-infection threshold while re-infection plays a major role above the re-infection threshold.

They also made provision in the model that if the effect of vaccination is increased the re-infection threshold also increases. There is however a disagreement as to whether the re-infection threshold exists [9].

Gabriela *et al.* [19] stated that if latent infection helps to sustain immunity to tuberculosis then treatment would increase the risk that subsequent infections lead to active TB disease. They also stated that the risk of reactivation may be higher in untreated individuals, who host a larger number of bacteria than those who are treated. Re-infection may then trigger reactivation of a latent infection, which suggests that treatment would reduce the risk of re-infection. Uys *et al.* [42] states that a decrease in mean delay time to diagnosis and a decrease in the progression rate from infection to active TB disease are two important control strategies for TB. The addition of TB treatment in models is an important factor to take into account.

Researchers have attempted to bring the effect of inherent genetic susceptibility to TB into their models. Murphy *et al.* [33] laid emphasis on how genetics can influence the effectiveness of treatment, while Singer *et al.* [38] focussed on re-infection.

The aspect of transmission is fundamental in the study of TB dynamics and the concept of cluster sizes is introduced by mathematical modellers. Aparicio *et al.* [3], brought the idea of cluster sizes into their model. They defined a cluster as the number of people an individual makes contact with daily such as members of that individual's household and workplace. An epidemiologically active cluster was defined as a cluster which has at least one individual with active TB disease. The reproduction number they considered is a function of cluster size. Murray [34], model predicts that the mean cluster size decreases as the prevalence of HIV increases despite the increase in the incidence of TB infection and active TB disease. This provides an indication that the incidence of extrapulmonary TB disease which is known as the non-infectious type of disease, increased. She also stated that the greater the proportion of people who are old aged, the higher the mean number of secondary infections produced by a single case. The converse should also hold. The concept of cluster sizes is especially valuable when considering interaction between different age groups.

Pešut *et al.* [35] points out that in developed countries tuberculosis is becoming an increasing problem among the elderly because of a natural decrease in immunity with age and

increasing longevity. Reactivation of latent TB disease which has been acquired earlier via transmission is the most important risk factor for the development of active TB disease [35, 45]. The risk of reactivation also increases with age since the population of latently infected individuals increases as age increases [47]. Since we also study a developed country in our comparison, it is important to know the basic TB dynamics of such a country.

2.2 TB in the Western Cape

HIV prevalence is certainly a factor which cannot be ignored when considering incidence of active TB disease since it is an important factor that drives the TB epidemic. The estimated HIV prevalence in the Western Cape is approximately 15 % [1, 25]. TB incidence varies with age. The incidence of active TB disease is about 20 times higher for an infant who is HIV infected than for one who is not HIV infected [21].

Wood *et al.* [50], stated that in 1990 Karel Styblo postulated that the impact of HIV infection is so large that the current tools available to control TB will be inadequate. The most commonly used diagnostic tools for TB is the TST. Its reliability has however been doubted. Despite the doubtful reliability for the sensitivity and specificity of the tuberculin skin test (TST), it is the most common immunological test for the estimation of prevalence, incidence and trend of Mtb infection in populations. Lawn *et al.* [26] found that the intense increase in TB notification rates during the period of 1996-2004 at a peri-urban township near Cape Town is associated with an increased number of HIV infected individuals. Currently the TB notification rates are the highest among the highly sexually active age classes whereas before the advent of the HIV epidemic it has been the case for those 60 years and older.

2.3 Age Structured Models

Continuous time models which incorporate age-structure in them are usually based on first order partial differential equations. Disease induced deaths are normally not included in mathematical models since it complicates the mathematical analysis of the models. Specifically, in TB models which conduct mathematical analysis, no distinction is made

between re-infection and reactivation. Whenever results are obtained for partial differential equations it is done with discrete methods such as finite difference methods or Markov chains.

Vynnycky *et al.* [45] discussed the change in re-infection rate of an individual who gets recurrently infected because of the change in immunity as depicted in FIG. 2.1. They assumed second or subsequent disease episodes within 5 years of initial infection or re-infection as being endogenous in nature if the individual has not been re-infected since the previous episode of active disease. Although provision was made for three separate active disease compartments of primary disease, endogenous reactivation and exogenous re-infection, it was assumed that the rate of recovery for these active disease compartments to the latent compartment is the same. Furthermore, the rate of proceeding from the infected and re-infected compartments to the latent compartment is assumed to be the same which is probably unrealistic since an individual's immunity may change upon re-infection.

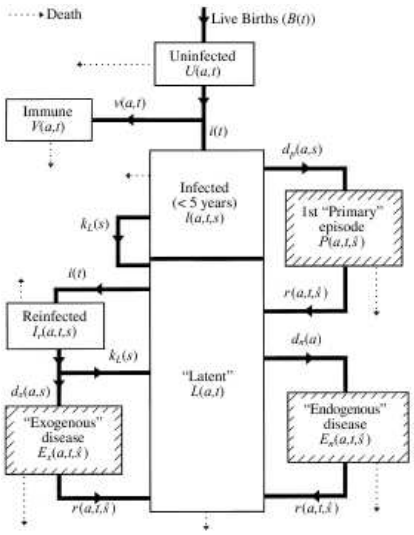


FIG. 2.1. Source: The model Vynnycky *et al.* [45] formulated took the effect of exogenous re-infection into account and differentiated between the 3 active disease compartments.

According to FIG. 2.2 different combinations of primary progression, exogenous re-infection and endogenous re-activation was fitted to active disease incidence for each age group. They also found that if exogenous re-infection is excluded from their model, the estimates for the risk of developing primary progression is unrealistically high.

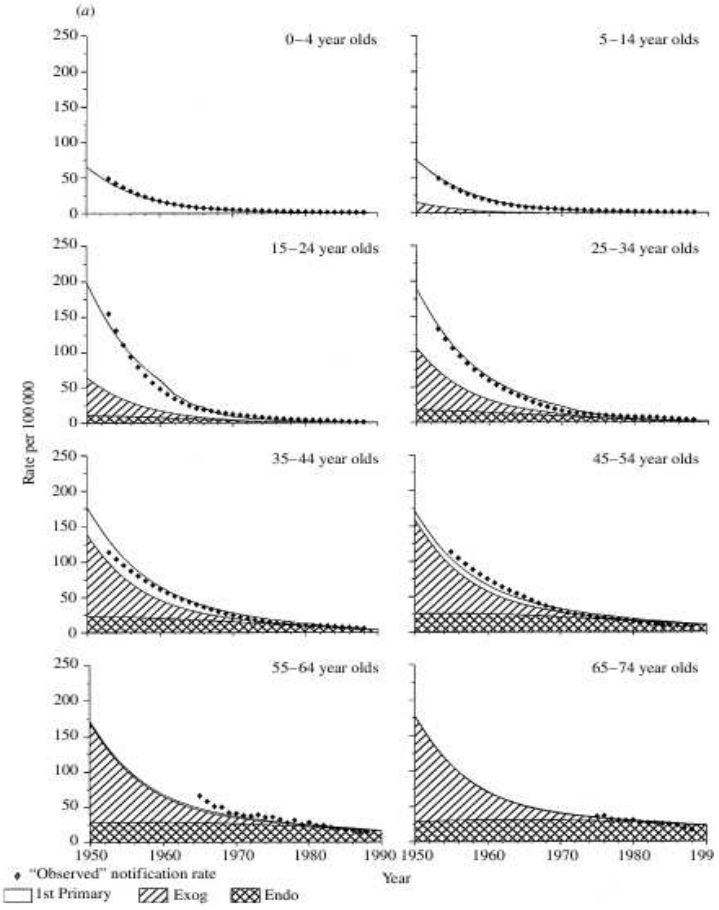


FIG. 2.2. Source: The importance of re-infection for each age group is illustrated by Vynnycky *et al.* [45].

The models of Uys *et al.* [41, 43] are modifications to the Vynnycky *et al.* [45] model as depicted in FIG. 2.3 and FIG. 2.4. They also differentiated between individuals who developed active TB disease by primary progression, endogenous re-activation and exogenous re-infection. Unlike in Vynnycky *et al.*, they assumed the rate of recovery for these three active disease compartments to the latent compartment to be different. An additional latent compartment was introduced to compensate for the absent effect of age in their model. A Markov model was used to find a relationship between incidence and re-infection rates.

Numerical methods were employed to simulate their model and no in depth mathematical analysis was done on the model.

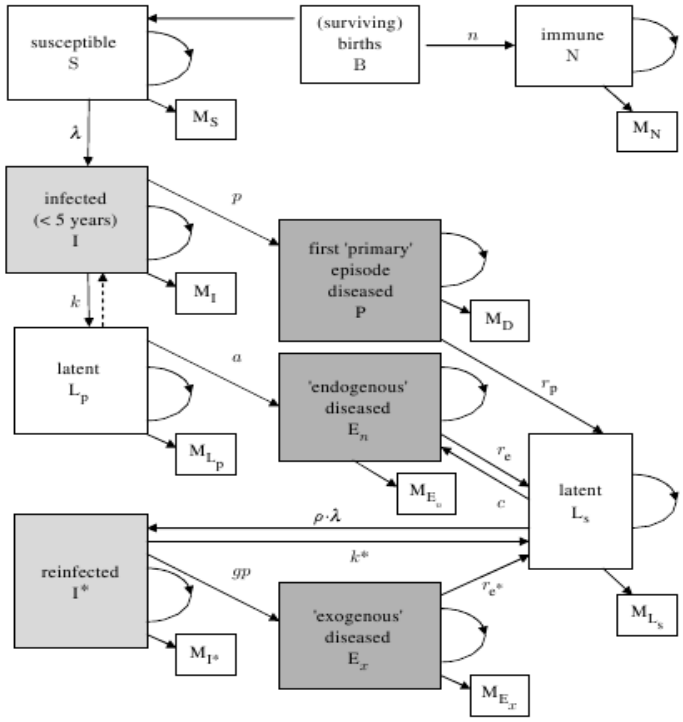


FIG. 2.3. Source: The modification of the Vynnycky *et al.* [45] model by Uys *et al.* [41].

Uys *et al.* [43] in turn differentiated between individuals who are drug-susceptible and those who are drug-resistant as shown in FIG. 2.4. The infection and re-infection rates are assumed to be the same which is possibly an unrealistic assumption since previous episodes of disease can determine the future episodes of disease within the same individuals.

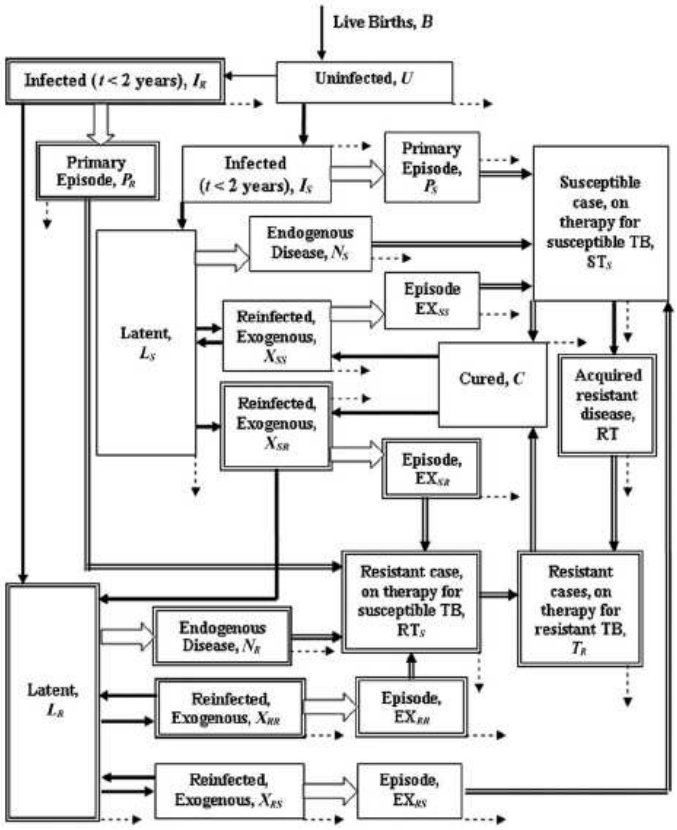


FIG. 2.4. Source: The modification of the Vynnycky *et al.* [45] model by Uys *et al.* [43]. A distinction is made between individuals who are drug-susceptible and those who are drug-resistant.

Castillo-Chavez *et al.* [12], introduced a susceptible-vaccinated-latent-infectious-removed model. They assumed partial efficiency of the vaccine. However, the effect of negative feedback in their model caused difficulties in the analysis of the model especially in proving the global stability result and the existence of an endemic steady states.

2.4 Formulation of the Model

There are two ways in which model fitting can be inaccurate. Firstly, if the fit is inaccurate and secondly, if the model is fitted to inaccurate data.

The finite difference scheme which Shim *et al.* [37] used for analysis of their model based

on first order partial differential equations, give rise to triangular matrices which do not provide a good fit to data. This difference scheme can obviously not be used, as the one half of the data will be fitted while the other half will be left void. The two point stencil scheme of Shim *et al.* [37] is considered but it is extended to a three point stencil scheme in this dissertation. Modifying their difference scheme get our heads around that problem.

The need for accurate data is very important as inaccurate data invalidates results. The lack of medical attention seeked by potentially active TB diseased individuals and problems associated with TB diagnosis, distorts the true reflection of the number of active TB disease cases within a country. Monney *et al.* [32] stated that immigrants to Switzerland tend to have active or latent tuberculosis. Illegal immigrants tend not to go for medical treatment since lack of identification may lead to deportation [17]. It is therefore possible that the considerable number of illegal immigrants in South Africa could lead to the underestimation of prevalence and incidence of active TB disease. Storla *et al.* [39] mentioned that TB is a rare disease and that almost all the patients who seek chronic cough treatment do not have TB. Such an assertion increases the diagnostic delay of TB patients. There is also a social stigma associated with TB that prevents people from getting tested. Directly observed therapy short course (DOTS) involves daily visits to TB clinics and patients get intimidated by the health care staff. This discourages TB patients from receiving full treatment and it also hampers the follow-up of patients.

2.5 Mathematical Preliminaries

2.5.1 Age Structures

We consider the McKendrik-von Förster partial differential equation as an example of an age-structured model for the illustration of the basic mathematical concepts.

Firstly, let $N(a,t)da$ = number of individuals of age $(a, a+da)$ at time t.

We consider the rate of change of the number of individuals in the age interval, $(a, a+da)$

as:

$$\frac{\partial}{\partial t}[N(a,t)\Delta a] = \text{rate of entry at age } a - \text{rate of exit at age } (a+\Delta a) - \text{natural deaths.}$$

This implies that

$$\frac{\partial N(a,t)}{\partial t}\Delta a = J(a,t) - J(a+\Delta a,t) - \mu(a)N(a,t)\Delta a,$$

where $\mu(a)$ is the per capita mortality rate for individuals aged a and $J(a,t)$ is the flux of individuals of age a at time t. The aging of individuals is portrayed in FIG. 2.5.

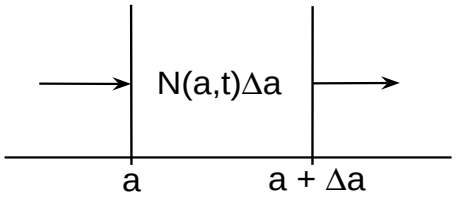


FIG. 2.5. The movement of individuals with respect to age.

We divide by Δa on both sides and obtain:

$$\frac{\partial N(a,t)}{\partial t} = -\left[\frac{J(a+\Delta a,t) - J(a,t)}{\Delta a}\right] - \mu(a)N(a,t)$$

Now lets take the limit $\lim_{\Delta a \to 0^+}$ on both sides:

$$\frac{\partial N(a,t)}{\partial t} = -\frac{\partial J(a,t)}{\partial a} - \mu(a)N(a,t)$$

Since all individuals grow older the flux is proportional to the density of the individuals. We can write the flux as:

$$J(a,t) = N(a,t)v(a,t),$$

with $v(a,t)$ being some speed of aging.

Moreover, since aging is only the passage of time we have: $v(a,t) = \frac{da}{dt} = 1.$

This leaves us with the following partial differential equation (PDE):

$$\frac{\partial N(a,t)}{\partial t} + \frac{\partial N(a,t)}{\partial a} = -\mu(a)N(a,t).$$

The birth or recruitment rate of the individuals does not form part of the PDE but it is taken as a boundary condition of the PDE:

$$N(0,t) = \int_0^{\omega} N(a,t)\mu(a)da,$$

where ω is a limiting age and the natural mortality rate and birth rate are assumed to be equal.

There is also an initial condition in the form of an initial age distribution: $N(a,0) = N_0(a)$.

2.5.2 The Method of Characteristics

Lets consider the partial differential equation (PDE):

$$\frac{\partial N(a,t)}{\partial t} + \frac{\partial N(a,t)}{\partial a} = -\mu(a)N(a,t).$$

The main idea of this method is to reduce a PDE into an ordinary differential equation

(ODE). This is done by imposing a change of variables so that

$$N(a,t) \quad = \quad N(\xi,\eta).$$

This change of variables will enable us to use the chain rule. Hence,

$$\frac{\partial N(a,t)}{\partial \xi} \quad = \quad \frac{\partial N(a,t)}{\partial a}\frac{\partial a(\xi,\eta)}{\partial \xi} + \frac{\partial N(a,t)}{\partial t}\frac{\partial t(\xi,\eta)}{\partial \xi} \quad = \quad -\mu(a)N(a,t)$$

with

$$\frac{\partial a(\xi,\eta)}{\partial \xi} \quad = \quad 1 \tag{2.1}$$

and

$$\frac{\partial t(\xi,\eta)}{\partial \xi} \quad = \quad 1, \tag{2.2}$$

which implies that

$$\frac{\partial N(a,t)}{\partial \xi} \quad = \quad -\mu(a)N(a,t). \tag{2.3}$$

Equation (2.3) of the above mentioned equations has still an a and a ξ. Since this is the case we must make a and ξ as close as possible. However, it is important that equations (2.1) and (2.2) hold. We thus need to use initial conditions for these differential equations. Firstly, it is required that a will be as close to ξ as possible so that $\xi = 0 \;\Rightarrow\; a = 0$. It is now important that t will be close to η so that $\xi = 0 \;\Rightarrow\; t = \eta$. By solving equations (2.1) and (2.2) we obtain:

$$a = \xi + c_1(\eta),$$

$$t = \xi + c_2(\eta).$$

Using the initial conditions $a(0,\eta) = 0$ and $t(0,\eta) = \eta$, we have

$$a = \xi$$

and

$$t = \xi + \eta \quad \text{or that} \quad \eta = t - a.$$

Thus equation (2.3) simplifies to

$$\frac{dN(\xi,\eta)}{d\xi} = -\mu(\xi)N(\xi,\eta).$$

By separation of variables we obtain:

$$\frac{dN(\xi,\eta)}{N(\xi,\eta)} = -\mu(\xi)d\xi. \tag{2.4}$$

The number in each cohort at some ξ (age) will depend on the boundary condition, $N(0,t)$ for $\eta \geq 0$ i.e. $a \leq t$ or on the initial condition, $N(a,0)$ for $\eta \leq 0$ i.e. $a \geq t$. So now equation (2.4) is solved separately for the regions, $a \leq t$ and $a \geq t$ as shown in FIG. 2.6.

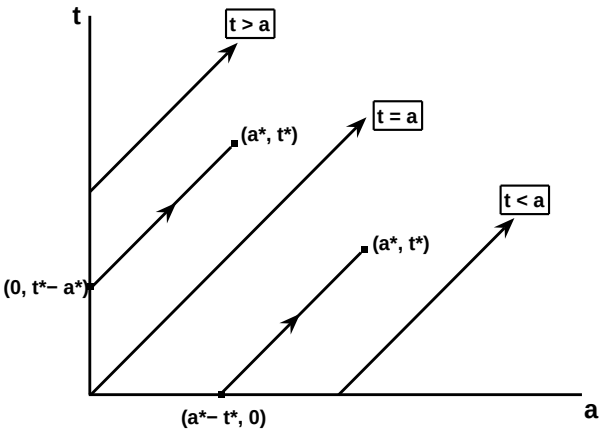


FIG. 2.6. As ξ gets increased it moves along a line of slope 1 and the line always start from the axes to the point (a^*, t^*) in the at plane.

For $a \leq t$ the boundary condition, $N(0, t-a)$ applies, so that

$$\begin{aligned}
\int_{\xi=0}^{\xi=a} \frac{dN(\xi)}{N(\xi)} &= -\int_{\xi=0}^{\xi=a} \mu(\xi) d\xi, \\
\Rightarrow [\ln N(\xi)]_{\xi=0}^{\xi=a} &= -\int_0^a \mu(\xi) d\xi, \\
\Rightarrow \ln N(a,t) - \ln N(0,t-a) &= -\int_0^a \mu(\xi) d\xi, \\
\Rightarrow \ln \left[\frac{N(a,t)}{N(0,t-a)}\right] &= -\int_0^a \mu(\xi) d\xi, \\
\Rightarrow N(a,t) &= N(0,t-a)\, e^{-\int_0^a \mu(\xi) d\xi}.
\end{aligned}$$

For $a \geq t$ the initial condition $N(a-t, 0)$ applies, so that

$$\begin{aligned}
\int_{\xi=a-t}^{\xi=a} \frac{dN(\xi)}{N(\xi)} &= -\int_{\xi=a-t}^{\xi=a} \mu(\xi) d\xi, \\
\Rightarrow [\ln N(\xi)]_{\xi=a-t}^{\xi=a} &= -\int_{a-t}^a \mu(\xi) d\xi, \\
\Rightarrow \ln N(a,t) - \ln N(a-t,0) &= -\int_{a-t}^a \mu(\xi) d\xi, \\
\Rightarrow \ln \left[\frac{N(a,t)}{N(a-t,0)}\right] &= -\int_{a-t}^a \mu(\xi) d\xi, \\
\Rightarrow N(a,t) &= N(a-t,0)\, e^{-\int_{a-t}^a \mu(\xi) d\xi}.
\end{aligned}$$

The final result we have is:

$$N(a,t) = \begin{cases} N(0,t-a)\ e^{-\int_0^a \mu(\xi)d\xi} & \text{if } a \leq t, \\ N(a-t,0)\ e^{-\int_{a-t}^a \mu(\xi)d\xi} & \text{if } a \geq t. \end{cases}$$

2.5.3 Sensitivity Analysis

Latin hypercube sampling (LHS) is a so-called stratified sampling without replacement technique, where the random parameter distributions are divided into N non-overlapping equiprobable intervals. Each parameter which we sample is guided by the specification of a probability density function (pdf) which depends on priori information. A uniform distribution is the obvious choice if no information is available on the parameter. If a parameter range is entirely biologically unrealistic the sensitivity analysis will yield meaningless results. However, these ranges are seldom known. We can only make parameter ranges very small if we have some priori knowledge available of them [31].

Blower *et al.* [7] mentioned that the number of simulations (sample size), N should be be at least more than the number of parameters analysed for sensitivity. If the interval of variation for some parameter is very large, the sampling can be performed on a log scale to prevent under-sampling in the outer ranges of the interval where the parameter values are assumed to be very small. Each interval of each parameter is sampled without replacement, in order for the whole range of each parameter to be explored. A LHS matrix is generated which consists of N rows which represent the number of simulations while k columns represent the number of varied parameters. We then simulate N model solutions by using each combination of parameter values or each row of the LHS matrix. A partial rank correlation coefficient (PRCC) provides a guide of the strength of a linear association between rank-transformed input data and rank-transformed output data after the linear effects on the output data of the remaining inputs are discounted [31].

2.6 Summary

In this chapter a short review of TB in the Western Cape is considered, an introduction is given to TB models, age structured models in general are examined and the matter of model fitting is discussed. There is also an additional section on mathematical preliminaries which describe mathematical tools which are closely related to the work done in this dissertation.

The effect of re-infection and reactivation on TB dynamics remains an open-ended question as is pointed out by Gabriela *et al.* [18, 20]. The parameters values which are approximated for this model are based on Lui *et al.* [28] and Cohen *et al.* [14]. The literature specifically based on age structured models has very limited mathematical analysis for models which are very similar to our model. It is clear that mathematical analysis of PDE's is considerably more complicated than that of ordinary differential equations. The finite difference scheme which Shim *et al.* [37] postulated in their study of rotavirus, certainly provided a sound platform for the finite difference scheme used in this dissertation. Finally, the approach of Latin Hypercube Sampling was discussed as a sensitivity analysis technique.

Chapter 3

An age structured TB model

3.1 Introduction

The incidence rates of active TB disease are highly age-dependent as was pointed out in Section 1.2. TB dynamics differ from country to country as regards age and time. We therefore choose to consider a TB model which includes age-structure because age is a very important factor in the dynamics of TB. We only consider pulmonary TB in this model since it is the only type of TB which contributes to TB transmission. We assume that the effective contact rate is a function of both age and time, while the other parameters are only age-dependent.

Our model will give rise to first order partial differential equations which are in terms of age and time. Numerical simulations are done based on these partial differential equations.

3.2 Model Description

We consider a five state deterministic model comprising of those who are susceptible to TB infection (S), those susceptibles who get vaccinated with BCG (V), individuals with latent TB infection (L), individuals with active TB disease (I), and those who recovered from active TB disease (R). We assume that

$$N = S + V + L + I + R,$$

where N is the total population size.

A proportion $\psi(a)$ where $\psi(a) \in [0,1]$ of susceptibles is vaccinated at an early stage in their lives. They move into the class of vaccinated individuals. If $\psi(a) = 0$, then no susceptible individuals get vaccinated and if $0 < \psi(a) \leq 1$, some individuals get vaccinated with $\psi(a) = 1$, indicating that all susceptibles are vaccinated. As the immunity of a vaccinated individual wanes, contact with an individual with active TB can result in infection. Upon infection these individuals progress to the compartment of the latently infected. We assume that vaccination reduces the risk of infection. The reduction in the risk of infection is measured by $\theta(a)$, where $\theta \in [0,1]$. If $\theta(a) = 0$ then the vaccinated individuals are immune and totally protected against infection. However if $\theta(a) = 1$, then all the vaccinated individuals are prone to infection. We assume that the vaccine wanes completely by the age of 15 years and according provision is made for that in the simulations. Depending on an individual's immunity, a susceptible individual can either move to the compartment of latently infected individuals if the pathogen is hypo-virulent or to the active disease compartment if the pathogen is hyper-virulent. A susceptible is considered subject to primary progression if he or she develops active TB disease within the first two years of infection. We assume that a proportion $p(a)$ of individuals progress slowly to latent TB upon infection. Therefore, if $p(a) = 0$ all susceptible individuals progress to active TB disease while if $p(a) = 1$ all susceptible individuals progress slowly to active TB disease. The susceptibles only move out of their state if they come in contact with infectious individuals in compartment, I.

Individuals in the latent compartment can move to the compartment of the active TB disease through two processes; firstly by exogenous re-infection and secondly endogenous reactivation. Latently infected individuals can be re-infected with active TB disease if they come in contact with infectious individuals. The parameter associated with the re-infection of latently infected individuals is $r_1(a)$. If $r_1(a) \in (0,1)$, the susceptibility decreases due to previous history of infection and if $r_1(a) > 1$, the susceptibility increases as a result of previous infections. Individuals with active TB can recover and move to the compartment of the recovered who can relapse and develop active disease again. The recovered cases can also get re-infected if they come in contact with infectious individuals. Similarly, $r_2(a)$ is the parameter linked with the re-infection of recovered individuals. If $r_2(a) \in (0,1)$, the susceptibility after recovery to re-infection decreases due to previous active TB development

and if $r_2(a) > 1$ the susceptibility after recovery to re-infection increases. The movement between compartments and the description of the classes and parameters (rates) are shown in FIG. 3.1 and TABLE. 3.1 respectively.

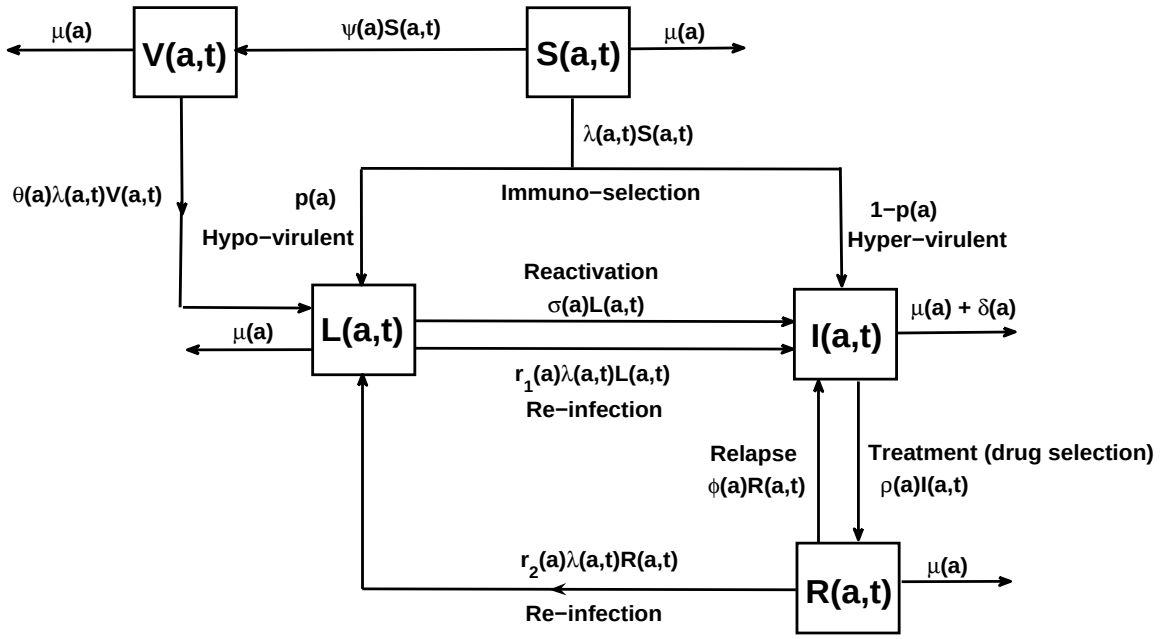


FIG. 3.1. SVLIR-Model which incorporates re-infection and reactivation of TB.

Class	Description
$S(a,t)$	Susceptibles – those who are at risk of getting the Mtb infection, (i.e. they are not infected or diseased).
$V(a,t)$	Vaccinated – those who are vaccinated at a very young age
$L(a,t)$	Latent class – those who are infected but not diseased
$I(a,t)$	Infectives – individuals with active disease and are infectious
$R(a,t)$	Recovered – those who are cured from the active disease either through treatment or self-recovery

Parameter	Description
$\mu(a)$	natural per capita death rate as a function of age
$\beta(a,t)$	effective contact rate for an individual of age a at time t
$\psi(a)$	rate of vaccination as a function of age
$\theta(a)$	proportion of individuals who have a reduced risk due to prior vaccination
$p(a)$	proportion of individuals who progress slowly to the latent phase
$r_1(a)$	re-infection parameter for the latent cases
$r_2(a)$	re-infection parameter related to susceptibility after recovery
$\rho(a)$	rate of recovery (by treatment or self-recovery)
$\phi(a)$	rate of relapse
$\delta(a)$	death rate related to TB
$\sigma(a)$	reactivation rate

TABLE. 3.1. Description of classes and parameters

Firstly, we let the force of infection be

$$\lambda(a,t) := \beta(a,t)I(a,t).$$

The law of mass action was used to set up the differential equations. The differential equations with respect to FIG. 3.1, combined with the model assumptions result in the following system of equations:

$$\frac{\partial S(a,t)}{\partial t} + \frac{\partial S(a,t)}{\partial a} = -[\lambda(a,t) + \mu(a) + \psi(a)]S(a,t), \tag{3.1}$$

$$\frac{\partial V(a,t)}{\partial t} + \frac{\partial V(a,t)}{\partial a} = \psi(a)S(a,t) - [\theta\lambda(a,t) + \mu(a)]V(a,t), \tag{3.2}$$

$$\frac{\partial L(a,t)}{\partial t} + \frac{\partial L(a,t)}{\partial a} = \lambda(a,t)[pS(a,t) + \theta V(a,t) + r_2 R(a,t)] - [r_1\lambda(a,t) + \sigma + \mu(a)]L(a,t), \tag{3.3}$$

$$\frac{\partial I(a,t)}{\partial t} + \frac{\partial I(a,t)}{\partial a} = \lambda(a,t)(1-p)S(a,t) + [\lambda(a,t)r_1 + \sigma]L(a,t) + \phi R(a,t) - [\mu(a) + \delta + \rho]I(a,t), \tag{3.4}$$

$$\frac{\partial R(a,t)}{\partial t} + \frac{\partial R(a,t)}{\partial a} = \rho I(a,t) - [r_2\lambda(a,t) + \phi + \mu(a)]R(a,t). \tag{3.5}$$

In this system the number of births and deaths are balanced, while the number of births

is equal to the number of people aged zero, $N(0,t) = S(0,t)$.

So the boundary conditions are

$$S(0,t) = \Pi(t) \quad \text{and} \quad V(0,t) = I(0,t) = L(0,t) = R(0,t) = 0.$$

By adding the boundary conditions and equations (3.1)-(3.5) respectively, we obtain

$$N(0,t) = \Pi(t) \quad \text{and}$$

$$\frac{\partial N(a,t)}{\partial t} + \frac{\partial N(a,t)}{\partial a} = -\mu(a)N(a,t) - \delta I(a,t).$$

If we take δ to be very small so that $\delta \approx 0$ then we have:

$$\frac{\partial N(a,t)}{\partial t} + \frac{\partial N(a,t)}{\partial a} = -\mu(a)N(a,t). \tag{3.6}$$

From Section 2.5.2 we have

$$N(a,t) = \begin{cases} \Pi(t)\ e^{-\int_0^a \mu(\xi)d\xi} & \text{if } a \leq t, \\ N(a-t,0)\ e^{-\int_{a-t}^a \mu(\xi)d\xi} & \text{if } a \geq t. \end{cases}$$

In order to non-dimensionalise the system of PDEs we introduce the following dimensionless variables

$$s(a,t) = \frac{S(a,t)}{N(a,t)}, \quad v(a,t) = \frac{V(a,t)}{N(a,t)}, \quad l(a,t) = \frac{L(a,t)}{N(a,t)}, \quad i(a,t) = \frac{I(a,t)}{N(a,t)} \quad \text{and} \quad r(a,t) = \frac{R(a,t)}{N(a,t)}.$$

Considering the class of the susceptibles, we have

$$\frac{\partial s(a,t)}{\partial t}+\frac{\partial s(a,t)}{\partial a} = \frac{\left[N(a,t)\frac{\partial S(a,t)}{\partial t}-S(a,t)\frac{\partial N(a,t)}{\partial t}\right]}{[N(a,t)]^2}+\frac{\left[N(a,t)\frac{\partial S(a,t)}{\partial a}-S(a,t)\frac{\partial N(a,t)}{\partial a}\right]}{[N(a,t)]^2}$$

$$= \frac{N(a,t)\left[\frac{\partial S(a,t)}{\partial a}+\frac{\partial S(a,t)}{\partial t}\right]}{[N(a,t)]^2} - \frac{S(a,t)\left[\frac{\partial N(a,t)}{\partial a}+\frac{\partial N(a,t)}{\partial t}\right]}{[N(a,t)]^2}.$$

Now by substituting equations (3.1) and (3.6) we obtain:

$$\frac{\partial s(a,t)}{\partial t}+\frac{\partial s(a,t)}{\partial a} = \frac{-[\lambda(a,t)+\mu(a)+\psi(a)]S(a,t)}{N(a,t)} - \frac{S(a,t)}{[N(a,t)]^2}[-\mu(a)N(a,t)].$$

$$\frac{\partial s(a,t)}{\partial t}+\frac{\partial s(a,t)}{\partial a} = -[\lambda(a,t)+\mu(a)+\psi(a)]s(a,t) + \mu(a)s(a,t).$$

Thus we have

$$\frac{\partial s(a,t)}{\partial t}+\frac{\partial s(a,t)}{\partial a} = -[\lambda(a,t)+\psi(a)]s(a,t). \tag{3.7}$$

Likewise, we can obtain PDEs for the non-dimensionalised classes $v(a,t)$, $l(a,t)$, $i(a,t)$ and $r(a,t)$. If we maintain the assumption $\delta \approx 0$, the equations respectively are;

$$\frac{\partial v(a,t)}{\partial t}+\frac{\partial v(a,t)}{\partial a} = \psi(a)s(a,t)-\lambda(a,t)v(a,t), \tag{3.8}$$

$$\frac{\partial l(a,t)}{\partial t}+\frac{\partial l(a,t)}{\partial a} = \lambda(a,t)[ps(a,t)+\theta v(a,t)+r_2\ r(a,t)] - (r_1\lambda(a,t)+\sigma)l(a,t), \tag{3.9}$$

$$\frac{\partial i(a,t)}{\partial t}+\frac{\partial i(a,t)}{\partial a} = \lambda(a,t)(1-p)s(a,t)+(\lambda(a,t)r_1+\sigma)l(a,t)+\phi r(a,t) - \rho i(a,t), \tag{3.10}$$

$$\frac{\partial r(a,t)}{\partial t}+\frac{\partial r(a,t)}{\partial a} = \rho i(a,t)-[r_2\lambda(a,t)+\phi]r(a,t). \tag{3.11}$$

Castillo-Chavez (1998) *et al.* [12] defined

$$\mathcal{V}_\psi(a) = e^{-\int_0^a \psi(u)du}$$

as the probability of susceptibles never receiving vaccination.

The disease free steady state is:

$$s(a) = \mathcal{V}_\psi(a), \quad v(a) = 1 - \mathcal{V}_\psi(a) \quad \text{and} \quad l(a) = i(a) = r(a) = 0.$$

In order to study the asymptotic behaviour of the non-dimensionalised system of equations we linearise equations (3.9) and (3.10). We consider exponential solutions which are of the form

$$l(a,t) = \mathcal{L}(a)e^{\gamma t}, \quad i(a,t) = \mathcal{I}(a)e^{\gamma t} \quad \text{and} \quad \lambda(t) = \lambda_0 e^{\gamma t} + O(e^{2\gamma t})$$

where

$$\lambda_0 = \int_0^\infty \mathcal{I}(a)\beta_\infty(a)da \tag{3.12}$$

with $\beta_\infty(a)$ being the transmission rate when the equilibrium is reached.

Equations (3.9) and (3.10) can now be written in the following form if t is assumed to be a parameter;

$$\gamma e^{\gamma t}\mathcal{L}(a) + e^{\gamma t}\frac{d\mathcal{L}(a)}{da} = \lambda_0 e^{\gamma t}\mathcal{F}_\psi(a) - (r_1\lambda_0 e^{\gamma t} + \sigma)e^{\gamma t}\mathcal{L}(a),$$

$$\Rightarrow \quad \gamma\mathcal{L}(a) + \frac{d\mathcal{L}(a)}{da} = \lambda_0\mathcal{F}_\psi(a) - (r_1\lambda_0 e^{\gamma t} + \sigma)\mathcal{L}(a), \tag{3.13}$$

and

$$\gamma e^{\gamma t}\mathcal{I}(a) + e^{\gamma t}\frac{d\mathcal{I}(a)}{da} = \lambda_0 e^{\gamma t}(1-p)\mathcal{V}_\psi(a) + (\lambda_0 e^{\gamma t} r_1 + \sigma)e^{\gamma t}\mathcal{L}(a) - \rho e^{\gamma t}\mathcal{I}(a),$$

$$\Rightarrow \quad \gamma\mathcal{I}(a) + \frac{d\mathcal{I}(a)}{da} = \lambda_0(1-p)\mathcal{V}_\psi(a) + (\lambda_0 e^{\gamma t} r_1 + \sigma)\mathcal{L}(a) - \rho\mathcal{I}(a), \tag{3.14}$$

with

$$\mathcal{F}_\psi(a) = p\mathcal{V}_\psi(a) + \theta[1 - \mathcal{V}_\psi(a)].$$

Firstly, considering equation (3.13) we have:

$$\gamma\mathcal{L}(a) + \frac{d\mathcal{L}(a)}{da} = \lambda_0\mathcal{F}_\psi(a) - (r_1\lambda_0 e^{\gamma t} + \sigma)\mathcal{L}(a),$$

$$\Rightarrow \quad (r_1\lambda_0 e^{\gamma t} + \sigma + \gamma)\mathcal{L}(a) + \frac{d\mathcal{L}(a)}{da} = \lambda_0\mathcal{F}_\psi(a).$$

By using integrating factor we obtain

$$\frac{d}{da}[\mathcal{L}(a)\exp[(r_1\lambda_0\exp(\gamma t) + \sigma + \gamma)a]] = \lambda_0\mathcal{F}_\psi(a)\exp[(r_1\lambda_0\exp(\gamma t) + \sigma + \gamma)a],$$

$$\Rightarrow \quad \mathcal{L}(a) = \exp[-(r_1\lambda_0\exp(\gamma t) + \sigma + \gamma)a]\times \int_0^a \lambda_0\mathcal{F}_\psi(\alpha)\exp[(r_1\lambda_0\exp(\gamma t) + \sigma + \gamma)\alpha]d\alpha. \tag{3.15}$$

Secondly we investigate (3.14) so that

$$\gamma\mathcal{I}(a) + \frac{d\mathcal{I}(a)}{da} = \lambda_0(1-p)\mathcal{V}_\psi(a) + (\lambda_0 e^{\gamma t} r_1 + \sigma)\mathcal{L}(a) - \rho\mathcal{I}(a),$$

$$\Rightarrow \quad (\gamma + \rho)\mathcal{I}(a) + \frac{d\mathcal{I}(a)}{da} = \lambda_0(1-p)\mathcal{V}_\psi(a) + (\lambda_0 e^{\gamma t} r_1 + \sigma)\mathcal{L}(a).$$

By using integrating factor again we obtain

$$\mathcal{I}(a) = \exp[-(\gamma+\rho)a]\times$$
$$\int_0^a [\lambda_0(1-p)\mathcal{V}_\psi(\alpha) + (\lambda_0\exp(\gamma t)r_1+\sigma)\mathcal{L}(\alpha)]\exp[(\gamma+\rho)\alpha]d\alpha. \tag{3.16}$$

Since (3.16) is so lengthy we break it up into parts so that for now we only consider

$$\exp[-(\gamma+\rho)a]\int_0^a (\lambda_0\exp(\gamma t)r_1+\sigma)\exp[(\gamma+\rho)\alpha]\mathcal{L}(\alpha)d\alpha. \tag{3.17}$$

Substitute (3.15) into (3.17) to obtain:

$$\exp[-(\gamma+\rho)a]\int_0^a (\lambda_0\exp(\gamma t)r_1+\sigma)\exp[(\gamma+\rho)\alpha]\exp[-(r_1\lambda_0\exp(\gamma t)+\sigma+\gamma)\alpha]\times$$
$$\left[\int_0^\alpha \lambda_0\mathcal{F}_\psi(b)\exp[(r_1\lambda_0\exp(\gamma t)+\sigma+\gamma)b]db\right]d\alpha. \tag{3.18}$$

In order to simplify (3.18) we apply partial integration.

$$\text{Let} \quad A = \int_0^\alpha \mathcal{F}_\psi(b)\exp[(r_1\lambda_0\exp(\gamma t)+\sigma+\gamma)b]db$$
$$\Rightarrow \quad dA = \mathcal{F}_\psi(\alpha)\exp[(r_1\lambda_0\exp(\gamma t)+\sigma+\gamma)\alpha]d\alpha.$$

$$dB = \lambda_0(\lambda_0\exp(\gamma t)r_1+\sigma)\exp[-(r_1\lambda_0\exp(\gamma t)+\sigma-\rho)\alpha]d\alpha$$
$$\Rightarrow \quad B = \frac{\lambda_0(\lambda_0\exp(\gamma t)r_1+\sigma)}{(\rho-r_1\lambda_0\exp(\gamma t)-\sigma)}\exp[-(r_1\lambda_0\exp(\gamma t)+\sigma-\rho)\alpha].$$

$$\exp[-(\gamma+\rho)a]\int_0^a (\lambda_0\exp(\gamma t)r_1+\sigma)\exp[(\gamma+\rho)\alpha]\mathcal{L}(\alpha)d\alpha$$
$$= \exp[-(\gamma+\rho)a]\left[\left[\frac{\lambda_0(\lambda_0\exp(\gamma t)r_1+\sigma)}{(\rho-r_1\lambda_0\exp(\gamma t)-\sigma)}\exp[-(r_1\lambda_0\exp(\gamma t)+\sigma-\rho)\alpha]\times\right.\right.$$

$$\int_0^{\alpha} \mathcal{F}_{\psi}(b)\exp[(r_1\lambda_0\exp(\gamma t)+\sigma+\gamma)b]db\Big]_{\alpha=0}^{\alpha=a}$$

$$-\int_0^a \frac{\mathcal{F}_{\psi}(\alpha)\lambda_0(\lambda_0\exp(\gamma t)r_1+\sigma)}{(\rho-r_1\lambda_0\exp(\gamma t)-\sigma)}\exp[(\gamma+\rho)\alpha]d\alpha\Bigg],$$

$$= \frac{\lambda_0(\lambda_0\exp(\gamma t)r_1+\sigma)}{(\rho-r_1\lambda_0\exp(\gamma t)-\sigma)}\exp[-(r_1\lambda_0\exp(\gamma t)+\sigma-\gamma)a]\int_0^a \mathcal{F}_{\psi}(b)\exp[(r_1\lambda_0\exp(\gamma t)+\sigma+\gamma)b]db$$

$$-\frac{\lambda_0(\lambda_0\exp(\gamma t)r_1+\sigma)}{(\rho-r_1\lambda_0\exp(\gamma t)-\sigma)}\int_0^a \mathcal{F}_{\psi}(\alpha)\exp[(\gamma+\rho)(\alpha-a)]d\alpha,$$

$$= \frac{\lambda_0(\lambda_0 e^{\gamma t}r_1+\sigma)}{(\rho-r_1\lambda_0 e^{\gamma t}-\sigma)}\int_0^a \mathcal{F}_{\psi}(\alpha)\Big[\exp[(r_1\lambda_0 e^{\gamma t}+\sigma)(\alpha-a)]e^{\gamma(\alpha+a)}-e^{(\gamma+\rho)(\alpha-a)}\Big]d\alpha.$$

From this result, expression (3.16) can be written as:

$$\mathcal{I}(a) = \frac{\lambda_0(\lambda_0 e^{\gamma t}r_1+\sigma)}{(\rho-r_1\lambda_0 e^{\gamma t}-\sigma)}\int_0^a \mathcal{F}_{\psi}(\alpha)\Big[\exp[(r_1\lambda_0 e^{\gamma t}+\sigma)(\alpha-a)]e^{\gamma(\alpha+a)}-e^{(\gamma+\rho)(\alpha-a)}\Big]d\alpha$$

$$+ \lambda_0(1-p)\int_0^a \mathcal{V}_{\psi}(\alpha)e^{(\gamma+\rho)(\alpha-a)}d\alpha. \tag{3.19}$$

Expression (3.19) for $\mathcal{I}(a)$ can be substituted into expression (3.12) to obtain:

$$\lambda_0 = \frac{\lambda_0(\lambda_0 e^{\gamma t}r_1+\sigma)}{(\rho-r_1\lambda_0 e^{\gamma t}-\sigma)}\int_0^{\infty}\int_0^a \mathcal{F}_{\psi}(\alpha)\Big[e^{(r_1\lambda_0 e^{\gamma t}+\sigma)(\alpha-a)}e^{\gamma(\alpha+a)}-e^{(\gamma+\rho)(\alpha-a)}\Big]\beta_{\infty}(\alpha+\tau)d\alpha da$$

$$+\lambda_0(1-p)\int_0^{\infty}\int_0^a \mathcal{V}_{\psi}(\alpha)e^{(\gamma+\rho)(\alpha-a)}\beta_{\infty}(\alpha+\tau)d\alpha da.$$

By changing the order of integration we let $\tau = a-\alpha$ and divide both sides by λ_0, we obtain

$$1 = \frac{\lambda_0 e^{\gamma t} r_1 + \sigma}{(\rho - r_1 \lambda_0 e^{\gamma t} - \sigma)} \int_0^\infty \int_0^\infty \mathcal{F}_\psi(\alpha) \left[e^{-(r_1 \lambda_0 e^{\gamma t} + \sigma)\tau} e^{\gamma(2\alpha+\tau)} - e^{-(\gamma+\rho)\tau} \right] \beta_\infty(\alpha+\tau) d\tau d\alpha$$
$$+ (1-p) \int_0^\infty \int_0^\infty \mathcal{V}_\psi(\alpha) e^{-(\gamma+\rho)\tau} \beta_\infty(\alpha+\tau) d\tau d\alpha. \tag{3.20}$$

The nett reproductive rate can be obtained, if we let $\gamma = 0$ for the right hand side of expression (3.20).

3.3 Numerical analysis and simulation of the model

Finite differences are used in the simulations since the software we use, does not make provision for first order PDEs. With these simulations we use a three-point stencil scheme since a stencil with too few points is not accurate while those with too many stencil points causes computational difficulties.

If PDEs get approximated with finite differences they need to get solved on a rectangular domain with $a \in [0, A]$ and $t \in [0, T]$ where A and T are the respective maximum age and maximum time under consideration. In this study backward differences are used. We divide this rectangular domain into a grid, as follows,

$$a_i = ih, \qquad (i = 1, ..., M)$$

$$t_j = jh, \qquad (j = 1, ..., N)$$

where h is the grid size so that $h = \frac{A}{M} = \frac{T}{N}$.

We approximate the derivatives by the analogy of Shim *et al.* [37]

$$\begin{aligned}\frac{\partial X(a,t)}{\partial a}+\frac{\partial X(a,t)}{\partial t} &\approx \frac{[X(a_i,t_j)-X(a_{i-1},t_j)]+[X(a_{i-1},t_j)-X(a_{i-1},t_{j-1})]}{h}\\ &= \frac{X(a_i,t_j)-X(a_{i-1},t_{j-1})}{h}\\ &= \frac{X_i^j-X_{i-1}^{j-1}}{h}\end{aligned}$$

where $X \in \{S,V,L,I,R\}$.

The discretised system of equations appears as follows:

$$\frac{S_i^j-S_{i-1}^{j-1}}{h} = -(\lambda_i^j+\mu_i+\psi_i)S_i^j, \tag{3.21}$$

$$\frac{V_i^j-V_{i-1}^{j-1}}{h} = \psi_i S_i^j-(\theta_i\lambda_i^j+\mu_i)V_i^j, \tag{3.22}$$

$$\frac{L_i^j-L_{i-1}^{j-1}}{h} = \lambda_i^j(p_iS_i^j+\theta_iV_i^j+r_{2i}R_i^j)-(r_{1i}\lambda_i^j+\sigma_i+\mu_i)L_i^j, \tag{3.23}$$

$$\frac{I_i^j-I_{i-1}^{j-1}}{h} = \lambda_i^j(1-p_i)S_i^j+(\lambda_i^jr_1+\sigma_i)L_i^j+\phi_iR_i^j-(\mu_i+\delta_i+\rho_i)I_i^j, \tag{3.24}$$

$$\frac{R_i^j-R_{i-1}^{j-1}}{h} = \rho_iI_i^j-(r_{2i}\lambda_i^j+\phi_i+\mu_i)R_i^j. \tag{3.25}$$

The integral for the force of infection get approximated by the trapezium rule as

$$\int_0^\infty \beta(a,t)I(a,t)da \approx \frac{1}{2}\beta_1^jI_1^jh + \sum_{i=2}^{i=M-1}\beta_i^jI_i^jh + \frac{1}{2}\beta_M^jI_M^jh.$$

By solving equations (3.21) - (3.25) we obtain

$$S_i^j = \frac{S_{i-1}^{j-1}}{1+h(\lambda_i^j+\mu_i+\psi_i)}, \tag{3.26}$$

$$V_i^j = \frac{h\psi_i S_i^j + V_{i-1}^{j-1}}{1+h(\theta_i\lambda_i^j+\mu_i)}, \tag{3.27}$$

$$L_i^j = \frac{\lambda_i^j h(p_i S_i^j + \theta_i V_i^j + r_{2i}R_i^j) + L_{i-1}^{j-1}}{1+h(r_{1i}\lambda_i^j+\sigma_i+\mu_i)}, \tag{3.28}$$

$$I_i^j = \frac{h[\lambda_i^j(1-p_i)S_i^j + (\lambda_i^j r_{1i}+\sigma_i)L_i^j + \phi_i R_i^j] + I_{i-1}^{j-1}}{1+h(\mu_i+\delta_i+\rho_i)}, \tag{3.29}$$

$$R_i^j = \frac{h\rho I_i^j + R_{i-1}^{j-1}}{1+h(r_{2i}\lambda_i^j+\phi_i+\mu_i)}. \tag{3.30}$$

The boundary conditions are approximated as

$$S_0^j = \Pi_j \quad \text{and} \quad V_0^j = I_0^j = L_0^j = R_0^j = 0.$$

Since the succeeding value of I_i^j is needed in all the difference equations, we thus use I_i^{j-1} as an approximation to I_i^j in the first calculation of our grids. We then use I_i^j again in the difference equations of the following calculations and run it a number of times until it converges to a solution.

Parameters for each age and year which corresponds to the modified data sets were obtained and then connected with each other by means of straight lines to form simple parameter functions.

The incidence of active TB disease is given by

$$\frac{(1-p)\beta(a,t)I(a,t)S(a,t)+\sigma L(a,t)+r_1\beta(a,t)I(a,t)L(a,t)+\phi R(a,t)}{S(a,t)+L(a,t)+R(a,t)}.$$

However, if the force of infection is taken as the integral it will only be a function of time and it will produce a tape-like three dimensional incidence function. This function is constant with respect to age but varies with respect to time.

So, it is therefore desirable to rather make the force of infection a function of both age and time. We can assume the following for the force of infection

$$\lambda(a,t) \ := \ \beta(a,t)I(a,t).$$

In discretised form it is written as

$$\lambda_i^j \ := \ \beta_i^j I_i^j.$$

In the search of the correct parameters we obtain the upper-triangular incidence matrix. Such a matrix is not useful in plotting numerical solutions on the $[0,A]\times[0,T]$ grid. We thus modify our finite difference scheme and approximate our derivatives by

$$\begin{aligned}\frac{\partial X(a,t)}{\partial a}+\frac{\partial X(a,t)}{\partial t} &\approx \frac{[X(a_i,t_j)-X(a_{i-1},t_j)]+[X(a_i,t_j)-X(a_i,t_{j-1})]}{h}\\ &= \frac{2X(a_i,t_j)-[X(a_{i-1},t_j)+X(a_i,t_{j-1})]}{h}\\ &= \frac{2X_i^j-(X_{i-1}^j+X_i^{j-1})}{h}\end{aligned}$$

where $X \in \ \{S,V,L,I,R\}$.

$$\frac{S_i^j - S_{i-1}^j}{h} + \frac{S_i^j - S_i^{j-1}}{h} = -(\lambda_i^j + \mu_i + \psi_i)S_i^j, \tag{3.31}$$

$$\frac{V_i^j - V_{i-1}^j}{h} + \frac{V_i^j - V_i^{j-1}}{h} = \psi_i S_i^j - (\theta_i \lambda_i^j + \mu_i)V_i^j, \tag{3.32}$$

$$\frac{L_i^j - L_{i-1}^j}{h} + \frac{L_i^j - L_i^{j-1}}{h} = \lambda_i^j(p_i S_i^j + \theta_i V_i^j + r_{2i}R_i^j) - (r_{1i}\lambda_i^j + \sigma_i + \mu_i)L_i^j, \tag{3.33}$$

$$\frac{I_i^j - I_{i-1}^j}{h} + \frac{I_i^j - I_i^{j-1}}{h} = \lambda_i^j(1-p_i)S_i^j + (\lambda_i^j r_{1i} + \sigma_i)L_i^j + \phi_i R_i^j - (\mu_i + \delta_i + \rho_i)I_i^j, \tag{3.34}$$

$$\frac{R_i^j - R_{i-1}^j}{h} + \frac{R_i^j - R_i^{j-1}}{h} = \rho_i I_i^j - (r_{2i}\lambda_i^j + \phi_i + \mu_i)R_i^j. \tag{3.35}$$

By solving equations (3.31) - (3.35) we obtain:

$$S_i^j = \frac{S_{i-1}^j + S_i^{j-1}}{2 + h(\lambda_i^j + \mu_i + \psi_i)}, \tag{3.36}$$

$$V_i^j = \frac{h\psi_i S_i^j + V_{i-1}^j + V_i^{j-1}}{2 + h(\theta_i \lambda_i^j + \mu_i)}, \tag{3.37}$$

$$L_i^j = \frac{\lambda_i^j h(p_i S_i^j + \theta_i V_i^j + r_{2i}R_i^j) + L_{i-1}^j + L_i^{j-1}}{2 + h(r_{1i}\lambda_i^j + \sigma_i + \mu_i)}, \tag{3.38}$$

$$I_i^j = \frac{h[\lambda_i^j(1-p_i)S_i^j + (\lambda_i^j r_{1i} + \sigma_i)L_i^j + \phi_i R_i^j] + I_{i-1}^j + I_i^{j-1}}{2 + h(\mu_i + \delta_i + \rho_i)}, \tag{3.39}$$

$$R_i^j = \frac{h\rho_i I_i^j + R_{i-1}^j + R_i^{j-1}}{2 + h(r_{2i}\lambda_i^j + \phi_i + \mu_i)}. \tag{3.40}$$

In order for us to prove the validity of the difference scheme, we need to establish the convergence of the numerical scheme.

We can let the error terms to be,

$$\zeta_i^j = S(a_i, t_j) - S_i^j, \tag{3.41}$$
$$\nu_i^j = V(a_i, t_j) - V_i^j, \tag{3.42}$$
$$\epsilon_i^j = L(a_i, t_j) - L_i^j, \tag{3.43}$$
$$\iota_i^j = I(a_i, t_j) - I_i^j, \tag{3.44}$$
$$\chi_i^j = R(a_i, t_j) - R_i^j. \tag{3.45}$$

Given that

$$\frac{2S_i^j - S_{i-1}^j - S_i^{j-1}}{h} = \frac{2S_i^j - 2S(a_i, t_j) - S_{i-1}^j + S(a_{i-1}, t_j) - S_i^{j-1} + S(a_i, t_{j-1})}{h} + \frac{2S(a_i, t_j) - S(a_{i-1}, t_j) - S(a_i, t_{j-1})}{h}. \tag{3.46}$$

We can write the expression in terms of the error term (3.41). Substituting the expression (3.41) into (3.46) to obtain

$$\frac{2S_i^j - S_{i-1}^j - S_i^{j-1}}{h} = -\frac{2\zeta_i^j - \zeta_{i-1}^j - \zeta_i^{j-1}}{h} + \frac{2S(a_i, t_j) - S(a_{i-1}, t_j) - S(a_i, t_{j-1})}{h}.$$

So,

$$\frac{2\zeta_i^j - \zeta_{i-1}^j - \zeta_i^{j-1}}{h} = -\frac{2S_i^j - S_{i-1}^j - S_i^{j-1}}{h} + \frac{2S(a_i, t_j) - S(a_{i-1}, t_j) - S(a_i, t_{j-1})}{h} \tag{3.47}$$

Substituting expression of (3.31) into (3.47) we have

$$\begin{aligned}\frac{2\zeta_i^j - \zeta_{i-1}^j - \zeta_i^{j-1}}{h} &= (\lambda_i^j + \mu_i + \psi_i)S_i^j \quad + \quad Q \\ &= (\lambda_i^j + \mu_i + \psi_i)[S_i^j - S(a_i, t_j)] \quad + \quad (\lambda_i^j + \mu_i + \psi_i)S(a_i, t_j) \quad + \quad Q \\ &= -(\lambda_i^j + \mu_i + \psi_i)\zeta_i^j \quad + \quad (\lambda_i^j + \mu_i + \psi_i)S(a_i, t_j) \quad + \quad Q\end{aligned}$$

where $Q = \dfrac{2S(a_i, t_j) - S(a_{i-1}, t_j) - S(a_i, t_{j-1})}{h}$.

If we write

$$O(h) = \lim_{h \to 0^+} \left[(\lambda_i^j + \mu_i + \psi_i)S(a_i, t_j) \quad + \quad \frac{2S(a_i, t_j) - S(a_{i-1}, t_j) - S(a_i, t_{j-1})}{h} \right],$$

we have

$$\frac{2\zeta_i^j - \zeta_{i-1}^j - \zeta_i^{j-1}}{h} = -(\lambda_i^j + \mu_i + \psi_i)\zeta_i^j \quad + \quad O(h). \tag{3.48}$$

Similarly, for the remaining equations we also have,

$$\frac{2\nu_i^j - \nu_{i-1}^j - \nu_i^{j-1}}{h} = \psi_i\zeta_i^j - (\theta_i\lambda_i^j + \mu_i)\nu_i^j \quad + \quad O(h), \tag{3.49}$$

$$\frac{2\epsilon_i^j - \epsilon_{i-1}^j - \epsilon_i^{j-1}}{h} = \lambda_i^j(p_i\zeta_i^j + \theta_i\nu_i^j + r_{2i}\chi_i^j) - (r_{1i}\lambda_i^j + \sigma_i + \mu_i)\epsilon_i^j \quad + \quad O(h), \tag{3.50}$$

$$\begin{aligned}\frac{2\iota_i^j - \iota_{i-1}^j - \iota_i^{j-1}}{h} &= \lambda_i^j(1 - p_i)\zeta_i^j + (\lambda_i^j r_{1i} + \sigma_i)\epsilon_i^j + \phi_i\chi_i^j \\ &\quad -(\mu_i + \delta_i + \rho_i)\iota_i^j \quad + \quad O(h),\end{aligned} \tag{3.51}$$

$$\frac{2\chi_i^j - \chi_{i-1}^j - \chi_i^{j-1}}{h} = \rho_i\iota_i^j - (r_{2i}\lambda_i^j + \phi_i + \mu_i)\chi_i^j \quad + \quad O(h). \tag{3.52}$$

By taking initial and boundary conditions of this system to be

$$\zeta_0^j = \nu_0^j = \epsilon_0^j = \iota_0^j = \chi_0^j = 0 \quad \text{for} \quad j = 1, 2, ..., N$$

and

$$\zeta_i^0 = \nu_i^0 = \epsilon_i^0 = \iota_i^0 = \chi_i^0 = 0 \quad \text{for} \quad i = 1, 2, ..., M, \text{ respectively,}$$

we can solve equations (3.48)–(3.52), to obtain solutions for ζ_i^j, ν_i^j, ϵ_i^j, ι_i^j, and χ_i^j.

$$\zeta_i^j = \frac{\zeta_{i-1}^j + \zeta_i^{j-1}}{2 + h(\lambda_i^j + \mu_i + \psi_i)} + O(h^2), \tag{3.53}$$

$$\nu_i^j = \frac{h\psi_i\zeta_i^j + \nu_{i-1}^j + \nu_i^{j-1}}{2 + h(\theta_i\lambda_i^j + \mu_i)} + O(h^2), \tag{3.54}$$

$$\epsilon_i^j = \frac{\lambda_i^j h(p_i\zeta_i^j + \theta_i\nu_i^j + r_{2i}\chi_i^j) + \epsilon_{i-1}^j + \epsilon_i^{j-1}}{2 + h(r_{1i}\lambda_i^j + \sigma_i + \mu_i)} + O(h^2), \tag{3.55}$$

$$\iota_i^j = \frac{h[\lambda_i^j(1-p_i)\zeta_i^j + (\lambda_i^j r_{1i} + \sigma_i)\epsilon_i^j + \phi_i\chi_i^j] + \iota_{i-1}^j + \iota_i^{j-1}}{2 + h(\mu_i + \delta_i + \rho_i)} + O(h^2), \tag{3.56}$$

$$\chi_i^j = \frac{h\rho_i\iota_i^j + \chi_{i-1}^j + \chi_i^{j-1}}{2 + h(r_{2i}\lambda_i^j + \phi_i + \mu_i)} + O(h^2). \tag{3.57}$$

It is important to prove that expressions (3.53)–(3.57) of the error terms converge to 0, as h approaches to 0. The proof of convergence validates the suitability of the difference scheme. We however deliberately avoid the proof because it is not central to the current work and usually convergence of numerical schemes is integrated into the numerical simulation codes in Matlab.

3.4 Summary

An age structured TB model is formulated in this Chapter. Some numerical analysis of the model is given. The model is discretised for the purpose of numerical simulations. In

the next chapter we now apply the discretised model by fitting it to the Taiwan data.

Chapter 4

Model fit to Taiwan data

We compare two data sets of active TB disease incidence rates in our study which are a data set from Yu *et al.* [51] for Taiwan and the Western Cape Tuberculosis Programme from the City of Cape Town Health Department. In this chapter we focus on fitting our model to the data from Taiwan.

The data set of active TB disease notification rates for Taiwan is obtained from Yu *et al.* [51] and is shown in TABLE. 4.1. The data is available in 5 age intervals. It is for the years 1957-1961, 1977-1981 and 1997-2001 applicable to 0-14, 15-24, 25-44, 45-64 and 65 or older, age intervals in years.

According to Yu *et al.* [51] data on the number of TB cases and patient age were collected from the National Tuberculosis Registry of Taiwan. In 1957, only bacteriologically confirmed pulmonary cases were registered. Cavitary pulmonary TB was included in the register from 1969. Cases with extensive parenchyma involvement were included in 1974, pleural TB cases in 1978, pathologically confirmed extrapulmonary TB cases in 1981, moderately advanced pulmonary TB cases in aboriginal and high incidence areas in 1984, and minimal pulmonary TB cases in aboriginal areas in 1988. Only since 1991 have all TB cases been included in the national register in Taiwan. The reliability of the TB notification rate is limited by completeness and accuracy of the reporting [51]. We shall assume for this study that the TB cases are predominately pulmonary TB cases. The other TB cases are assumed to be negligible.

The data set is significant since it can serve as a baseline for the study. We assume that the impact of HIV can be neglected in this data set. It is certainly related to our study since the data was collected over time and different age groups are considered.

		Age					
		0-14	**15-24**	**25-44**	**45-64**	**≥ 65**	**Unknown**
Year	**1957**	22	300	1380	668	57	30
	1958	29	681	2411	1289	133	26
	1959	44	684	3699	2191	313	81
	1960	43	829	2909	1877	268	49
	1961	59	766	2846	1961	282	73
	1977	55	1013	1666	3138	963	20
	1978	83	1015	1583	3174	971	25
	1979	54	831	1296	2736	586	28
	1980	41	704	1295	2384	1342	63
	1981	66	837	1436	2645	1264	49
	1997	170	1050	3389	668	4382	—
	1998	129	1031	2957	4029	6023	—
	1999	129	979	2726	3759	5903	—
	2000	141	1026	2799	3752	6192	—
	2001	139	918	2847	3880	6702	—

TABLE. 4.1. Source: Yu *et al.* [51], Notification rates for active TB disease of Taiwan from 1957-2001 divided in 5 age intervals for 3 periods of five consecutive years.

We assume that there is a uniform distribution of notification rates within the age intervals. The data of an age interval is divided by the length of the interval to obtain the average notification rate for that interval. We then assume that the average notification rate applies to the midpoint of the interval. The oldest age interval of the Taiwan data set is assumed to be 20 years long. The *unknown* column of TABLE. 4.1 represents the individuals with active TB disease whose ages are not known. This column is divided proportionally between the different age intervals according to the contribution of each age interval to the total notifications for that year which belong to a specific age interval. The formula of the modification for individuals in a specific age group with active TB disease is

$$\text{Estimated number of cases for the midpoint of age interval} \;=\; \frac{\mathcal{A} \;+\; \mathcal{B} \times \frac{\mathcal{A}}{\mathcal{C}}}{\mathcal{D}},$$

where

$\mathcal{A}$ = Cases for a specific age group within a specific year,
$\mathcal{B}$ = Cases from an unknown age group for a specific year,
$\mathcal{C}$ = Sum of cases from known age classes for a specific year,
$\mathcal{D}$ = Length of the specific age interval under consideration.

The modified data for Taiwan is displayed in TABLE. 4.2.

		Age				
		7	**20**	**35**	**55**	**75**
	1957	1.485	30.371	69.853	33.813	2.885
	1958	1.944	68.49	121.24	64.819	6.688
	1959	2.968	69.199	187.111	110.83	15.833
	1960	2.89	83.585	146.653	94.626	13.511
	1961	3.982	77.546	144.056	99.26	14.274
Year	**1977**	3.677	101.596	83.544	157.359	48.291
	1978	5.554	101.872	79.44	159.281	48.728
	1979	3.618	83.523	65.13	137.496	29.449
	1980	2.763	71.169	65.457	120.502	67.833
	1981	4.435	84.356	72.363	133.287	63.696
	1997	11.333	105	169.45	219.1	319.75
	1998	8.6	103.1	147.85	201.45	301.15
	1999	8.6	97.9	136.3	187.95	295.15
	2000	9.4	102.6	139.95	187.6	309.6
	2001	9.267	91.8	142.35	194	335.1

TABLE. 4.2. Modified notification rates for active TB disease of Taiwan from 1957-2001.

We assume the same set of initial conditions as that of Liu *et al.* [28] which is based on the TB dynamics of China. We changed the initial conditions of China proportionally to adjust for the population size in Taiwan and we also divided the value of the initial condition between different ages. We considered the year 1950 to be the initial year. We also assume that most of the parameters for Taiwan are similar to Liu *et al.* [28].

The boundary condition for the Taiwan susceptibles is

$$S(0,t) = 247093\exp[0.0145(t-1950)].$$

The recruitment rate is assumed to increase exponentially with time as the Taiwanese population grew at more or less 1% per annum [23, 40]. We assume that the recruitment rate increases in accordance with the increase in population over time.

The initial conditions of the classes are also specified.

The susceptibles' initial condition is

$$S(a,0) = \frac{247238}{1+0.0185a^2}.$$

It is a decreasing function with age since once you are infected you do not return back to the susceptibles compartment.

The vaccinated's initial condition is

$$V(a,0) = \frac{210611}{1+0.0185a^2}.$$

Likewise, vaccinated individuals' initial condition is a decreasing function with age since once you are infected you do not return back to the vaccinated compartment.

The latently infected individuals' initial condition is

$$L(a,0) = 19017[\exp(0.02a)-1].$$

It is an increasing function of age. The reason for this is once you are infected you do not completely get cured from TB so one can expect that the number of latently infected individuals will only rise as age increases.

The infectious individuals' initial condition is

$$I(a,0) = 2.9058a(100-a).$$

It is also an increasing function of age but then decreases at about the age of 50 when individuals start to die from aging. As more infectious individuals die from aging there are

fewer infectious individuals who remain alive.

The removed individuals' initial condition is

$$R(a,0) \quad = \quad 5.4123a(100-a).$$

The number of removed individuals is an increasing function of age but then decreases at about the age of 50 when individuals also start to die from aging. As more removed individuals die from aging there are fewer removed individuals who remain alive.

The boundary conditions are calculated by multiplying the population of Taiwan in 1950 by the proportion of individuals up to 10 years and then divide by 10 which is the length of the interval. We can do this if we assume that there is a uniform distribution of individuals within the age interval. In this way, we obtain an estimate for the number of individuals who are in the first year of their lives, in other words an estimate for the recruitment rate is obtained. The boundary condition is also made an increasing function of time in terms of the average annual growth rate of the population [15].The assumptions for the boundary condition as a function of time and the initial conditions as a function of age are depicted in FIG. 4.1.

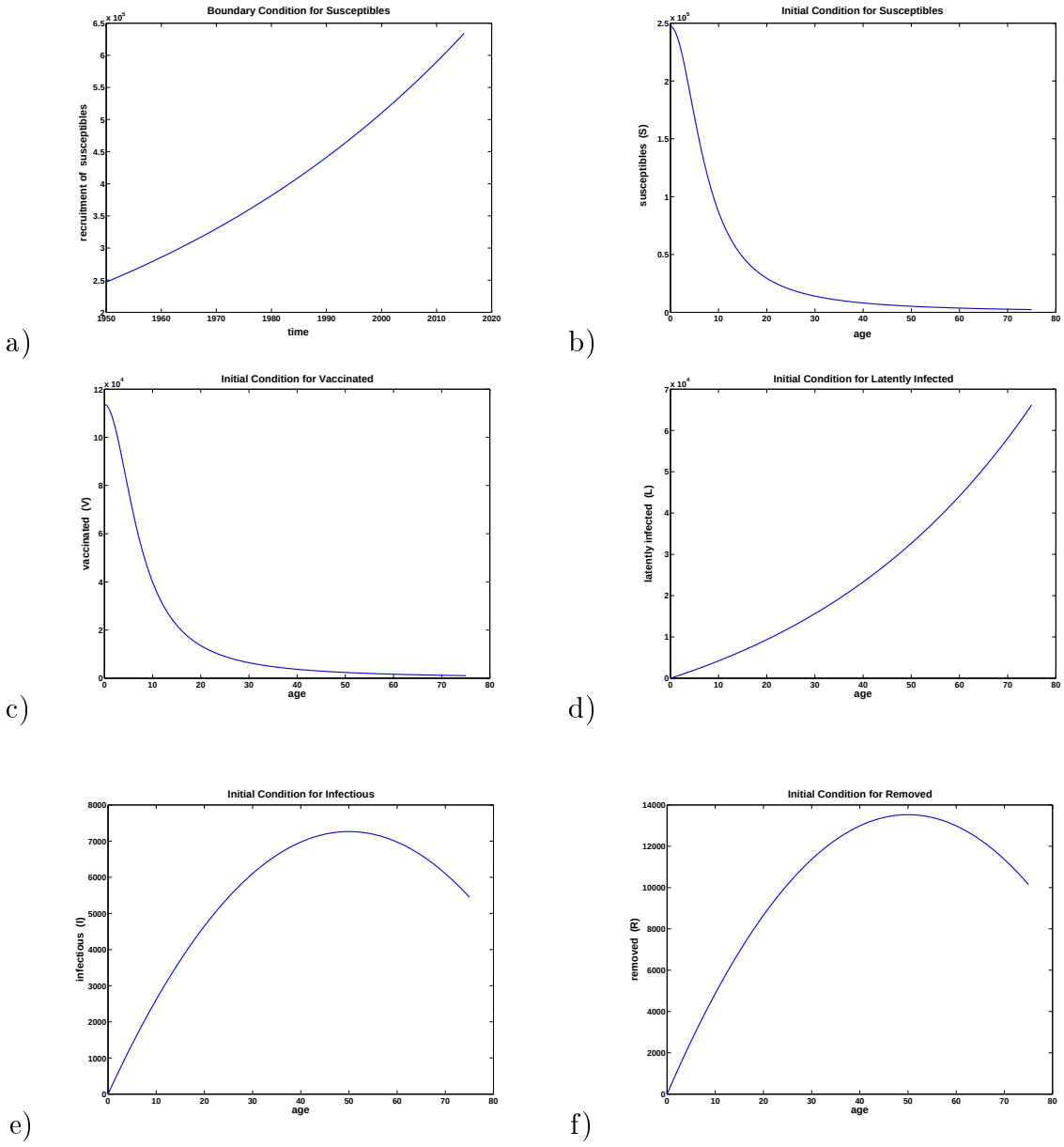


FIG. 4.1. (a) show how the recruitment of susceptibles evolved over time. We have called this a boundary condition $S(0,t)$. The graph is found by plotting the function $S(0,t) = 247093\exp[0.0145(t-1950)]$. The initial conditions are: (b) $S(a,0) = \dfrac{247238}{1+0.0185a^2}$ for the susceptibles, (c) $V(a,0) = \dfrac{210611}{1+0.0185a^2}$ for the vaccinated, (d) $L(a,0) = 19017[\exp(0.02a)-1]$ for the latently infected, (e) $I(a,0) = 2.9058a(100-a)$ for the infectious and (f) $R(a,0) = 5.4123a(100-a)$ for the removed individuals.

The natural mortality rate is approximated by a Gompertz function which is based on Ukraine mortality data investigated by Brooks-Pollock *et al.* [11]. The same mortality function is assumed in this study. The values of the parameters in TABLE. 4.3 for various age groups are either assumed or approximated from the literature.

	Age					
Parameter	**At Birth**	**7**	**20**	**35**	**55**	**75**
μ	$0.0003\exp(0.0735 \times a)$ [11]					
ψ	0.750	0.08651251	0	0	0	0
p	0.8*	0.9*	0.88*	0.92*	0.93*	0.95*
r_1	0.0045	0.00325	0.45	1.45	1.75	1.45
r_2	0.0045	0.00090510	0.0265	0.1365	0.14	0.1
ρ	0.75**	0.975**	0.875**	0.9675**	0.575**	0.6**
ϕ	0.02**	0.875**	0.0175**	0.065**	0.1065**	0.07065**
δ	0.007***	0.00071***	0.0015125***	0.0025125***	0.005138***	0.015***
σ	5×10^{-9}**	5×10^{-8}**	0.000002**	0.00022**	0.00028**	0.0004**
θ	0.05	0.45	0.55	0.65	0.75	0.95

TABLE. 4.3. Parameter values for different ages. Sources: All the parameter values with * are approximated from Feng *et al.* [17], the parameter values with ** are approximated from Cohen *et al.* [14] and those with *** are approximated from Lui *et al.* [28]. The natural mortality rate is obtained from Brooks-Pollock *et al.* [11]. The other parameters are based on assumptions.

We fitted the model to the data by changing the $\beta-$parameter the effective contact rate as shown in TABLE. 4.4. The $\beta-$parameters for the "golden ages" are so high since it is appropriate to the first 15 year age interval which includes high incidence of active TB disease amongst infants.

	Age				
	At Birth	**3**	**7**	**13**	**20**
Maximum	0.0013	0.0064	0.04	0.0204	0.0234
Minimum	0.00001	0.00001	0.00001	0.00001	0.0002
Average	0.0003	0.0018	0.0243	0.0119	0.0130

	Age					
	27	**35**	**45**	**55**	**65**	**75**
Maximum	0.0219	0.028	0.0253	0.0293	0.0512	0.2055
Minimum	0.0015	0.0026	0.0031	0.0019	0.0009	0.0001
Average	0.0077	0.0081	0.0101	0.0154	0.0257	0.0765

TABLE. 4.4. Effective contact rates (β's) regarding the Taiwan data set for different ages.

4.1 Sensitivity Analysis

Sensitivity analysis is done in order to investigate how the model output changes if the input parameters are changed. Latin Hypercube Sampling and Partial Rank Correlation Coefficients (PRCCs) were used with a 1000 simulations per run. The magnitude of the absolute value of the PRCCs indicates the significance of the input parameter as regards to their sensitivity on model outputs. The removed individuals of age 75 years were chosen in our sensitivity analysis as portrayed in FIG. 4.2. The results are depicted in FIG. 4.2. This analysis was extended to all the age groups for different parameter values that were age specific. The results were however the same. We thus focus on the analysis presented here as representative of the sensitivity analysis of the model.

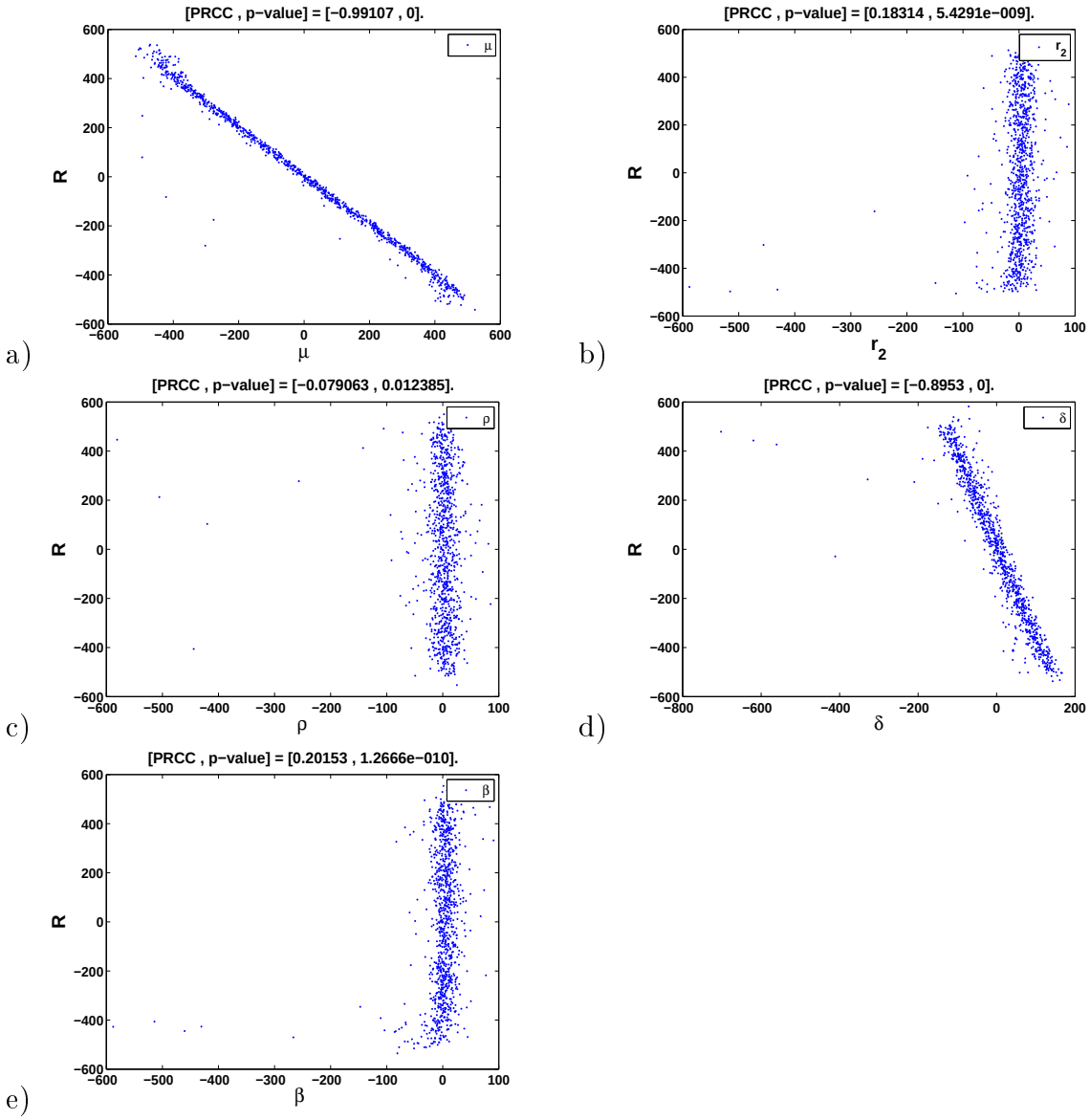


FIG. 4.2. PRCC scatter plots at time 600 years of the most significant parameters which are (a) μ, (b) r_2, (c) ρ, (d) δ and (e) β. The residuals of the linear regression between the rank-transformed values of the output of the removed individuals against rank-transformed values of the significant parameters.

FIG. 4.3 represents the bar plot of the PRCCs related to the different parameters. The lengths of the different bars indicate the magnitude in sensitivity of the number of removed individuals when parameters are changed.

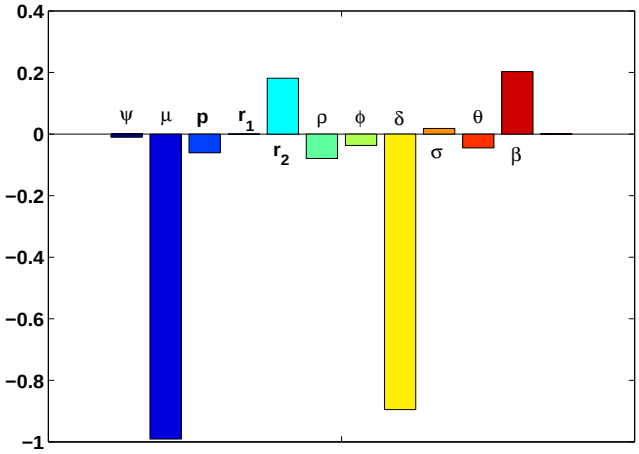


FIG. 4.3. PRCC bar plot of the input parameters with sample size 1000. The parameters with the most significant PRCCs are μ, r_2, ρ, δ and β.

If we take $h = 0.5$ and only assign values for the $\beta's$ of the ages and years corresponding to the data points, with the $\beta-$values being linearly connected to each other, the simulation results do not reflect or fit to the data. Incidence of active TB disease is sensitive to the changes in δ and μ which implies that they could be changed a little and produce different and realistic results. However, the incidence of active TB disease is slightly sensitive to the change in input parameters r_2, ρ and β. This analysis can be extended to the Cape Town Metropole data and the results are presented in the next chapter.

The data of TABLE. 4.2 is illustrated in FIG. 4.4 by means of a three dimensional plot. The figure shows the distribution of incidence over age and time. It is important to note the high incidence values for the old aged and how the incidence has steadily increased over time for some age groups.

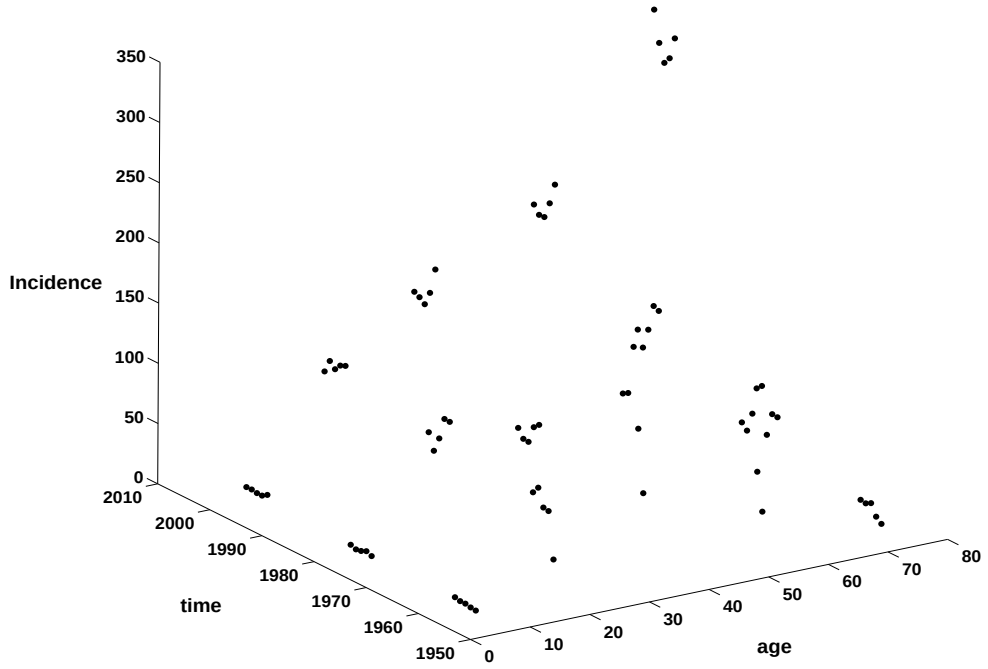


FIG. 4.4. Taiwan data plotted in three dimensional space with respect to time and age.

The model is thus fitted to the data in FIG. 4.4. A surface plot is obtained. Not only does this show a smooth data distribution but it also provides estimates for the incidence for the missing data.

The simulations are considered from different view angles as shown in Figures (4.5) to (4.8) to obtain a sound insight in the age-specific incidence for Taiwan. For the interpretations of colour in the figures the reader is referred to the electronic copy of the thesis. We now use the discretised model (3.31)-(3.35) and fit it to the data plotted in FIG. 4.4. By an iterative process, that involves changing of parameter values for each age group we obtain the mesh fit depicted in FIG. 4.5. The figure has a corresponding colour bar to show how the incidence changed over years. The model fit allows us to estimate the incidence of TB during the years 1962-1976 and 1982-1996 in which no data is available.

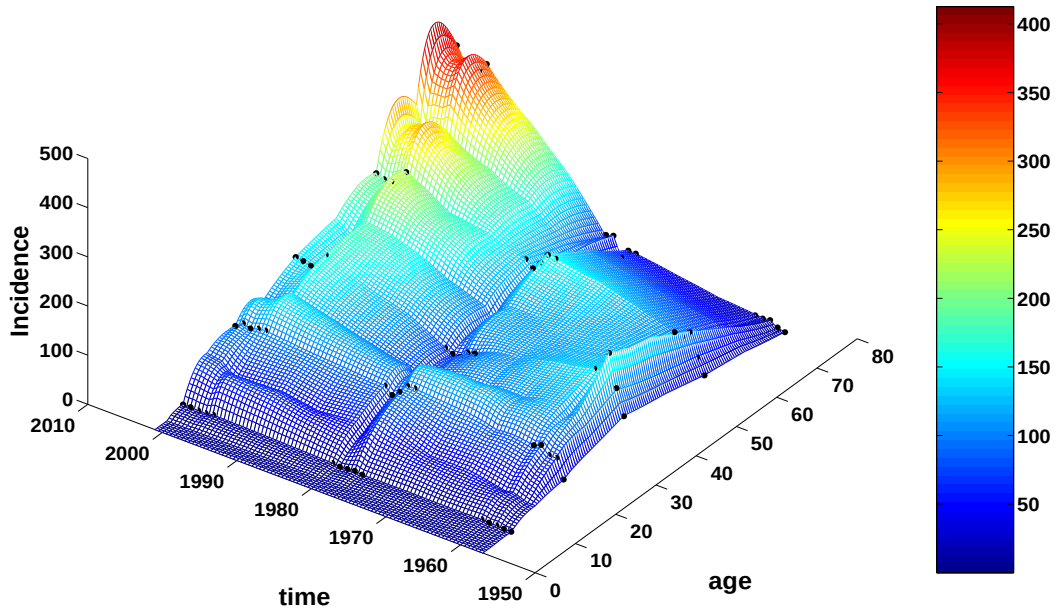


FIG. 4.5. A general view of the simulation together with the data points.

While FIG. 4.5 is clear and easily interpretable, it may be interesting to look at the same figure from different views. If we rotate the three dimensional FIG. 4.5 in such a way that incidence and age appear on the vertical axes, we can clearly observe how the incidence changes over time and age. This is specifically observed by looking at the colour bar. A steady increase in TB incidence is clearly observed with the biggest increase being observed in the old aged.

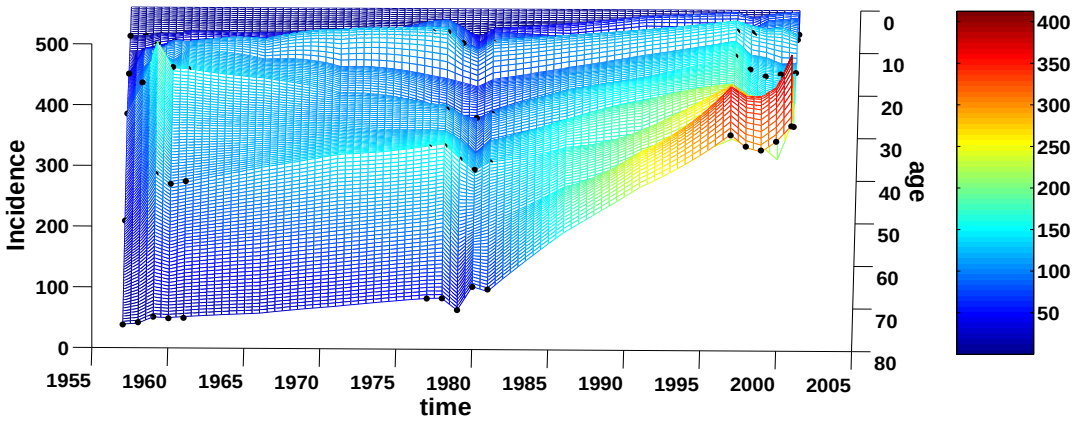


FIG. 4.6. A crude view of how the incidence for active TB disease changes over time.

A view from the top and the bottom of the three dimensional diagram FIG. 4.5 is shown in figures 4.7 and 4.8. This is because some data points appear on top of the mesh during fitting while others appear on bottom of the mesh during fitting. It is thus important to have these two views to determine outliers on our data.

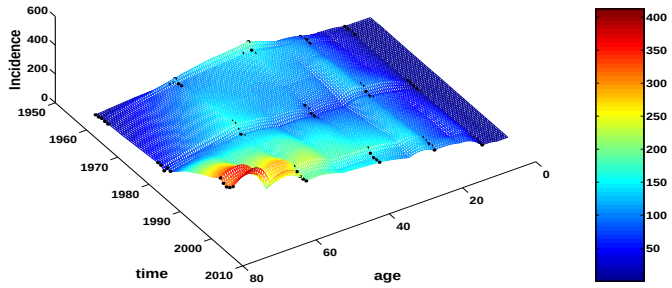


FIG. 4.7. A crude view of how the incidence for active TB disease changes over age.

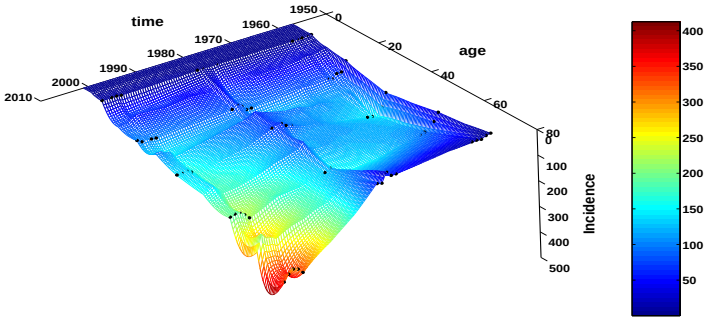


FIG. 4.8. Best fit of the model to the Taiwan data, which is considered from below.

If we focus on the change in incidence with age, we have the plot FIG. 4.9 which clearly show that TB incidence in Taiwan has been increasing with age over the years.

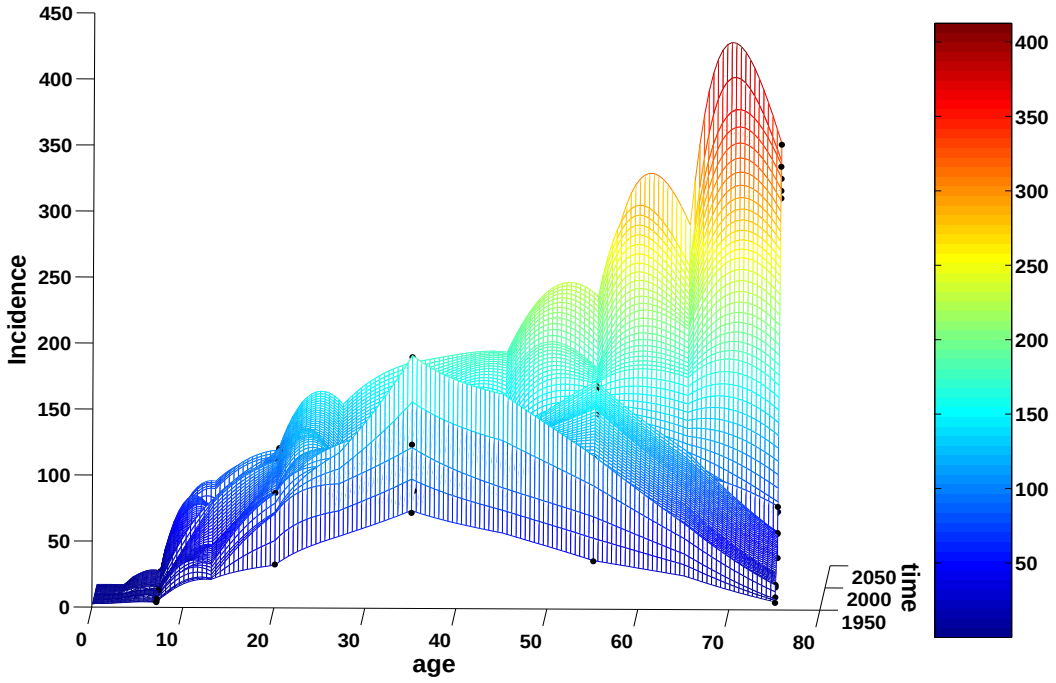


FIG. 4.9. Best fit of the model to the Taiwan data, which is considered along ages.

The change in incidence over time from 1957 to 2001 is clearly depicted in FIG. 4.10. The figures show that in all age groups TB incidence steadily increased although it is more remarkable in the old ages. This was also observed in FIG. 4.6.

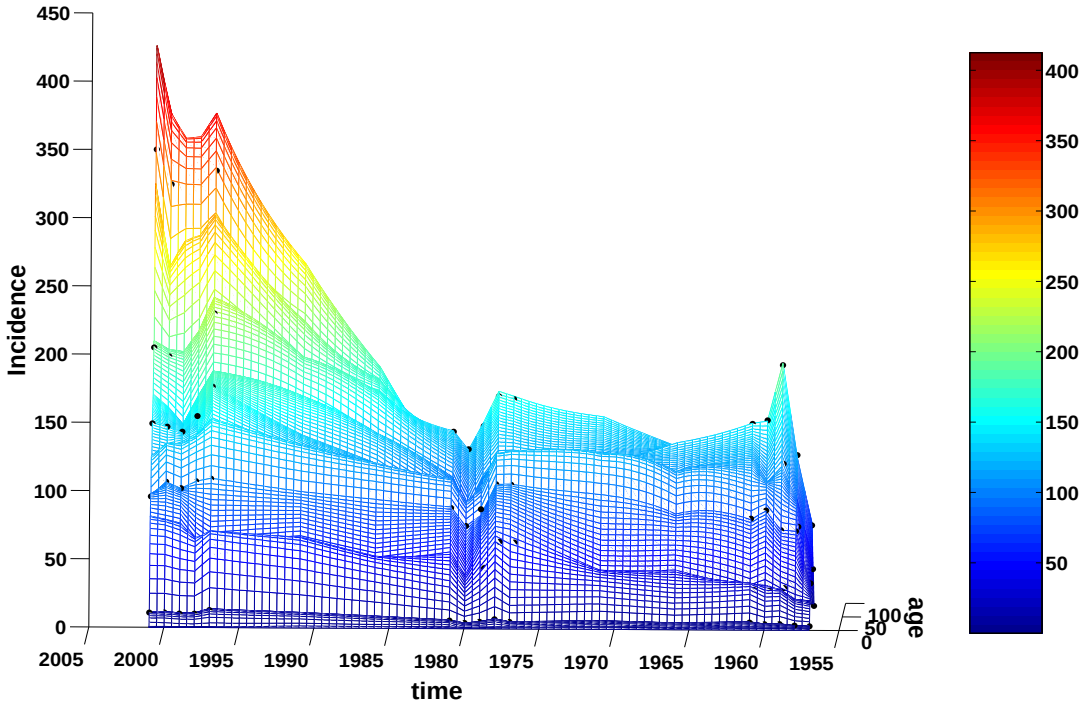


FIG. 4.10. Best fit of the model to the Taiwan data, which is considered along time in years.

Figures 4.7 to 4.10 can be summarised into a simple and very informative mesh colour map shown in FIG. 4.11. The incidence can be interpreted from the adjacent colour bar. It is clear that the incidence is high in the old aged.

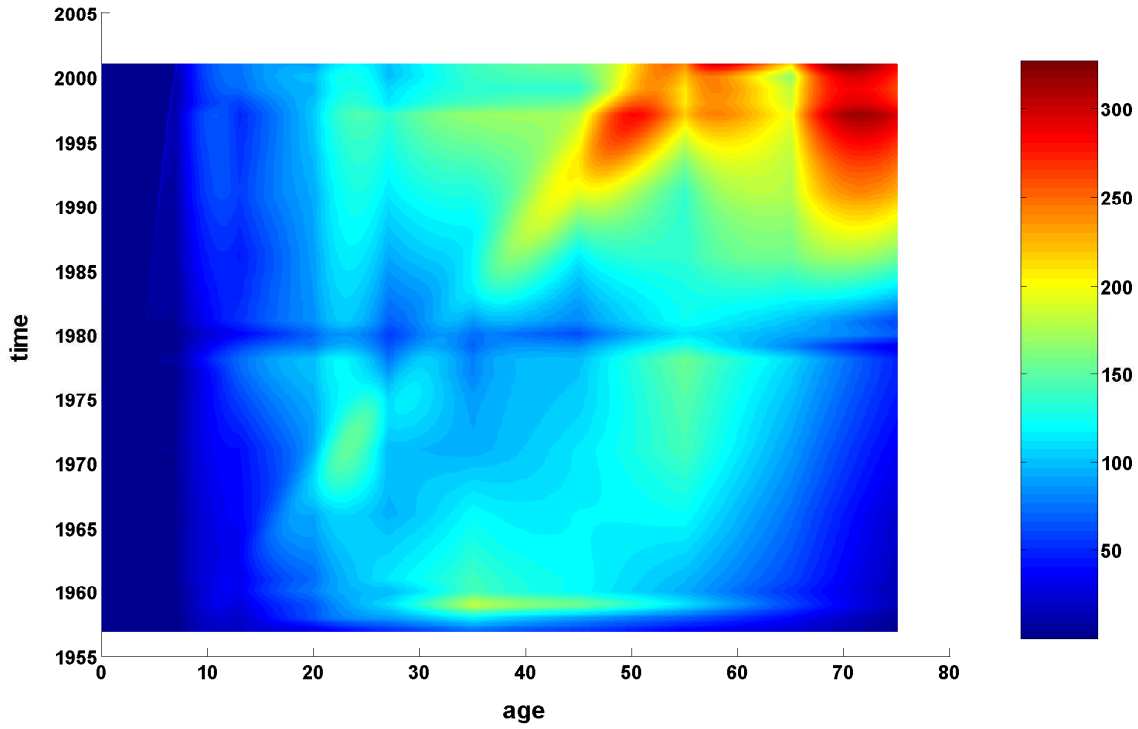


FIG. 4.11. Best fit of the model to the Taiwan data for the 5 age groups from 1957-2001, considered as a mesh colour map.

4.2 Projections

Projection in the case of active TB disease incidence is the process of making statements about the past and future incidence whose actual outcomes have not yet been observed based on the incidence which has been observed. Projections are useful in the sense that they give insights in the time trends which provide policymakers with a guide as to how they can approach the future progression of the disease.

The simulations of incidence over time for each age group are illustrated in FIG. 4.12. From FIG. 4.12 the incidence rates for children in the 0-14 year age interval and for individuals in the 15-24 year, 25-44 year and 45-64 year age intervals increase slightly over time and therefore the projection is also slightly increasing. The 65 years and older individuals exhibit a sharp increase in incidence rates over time.

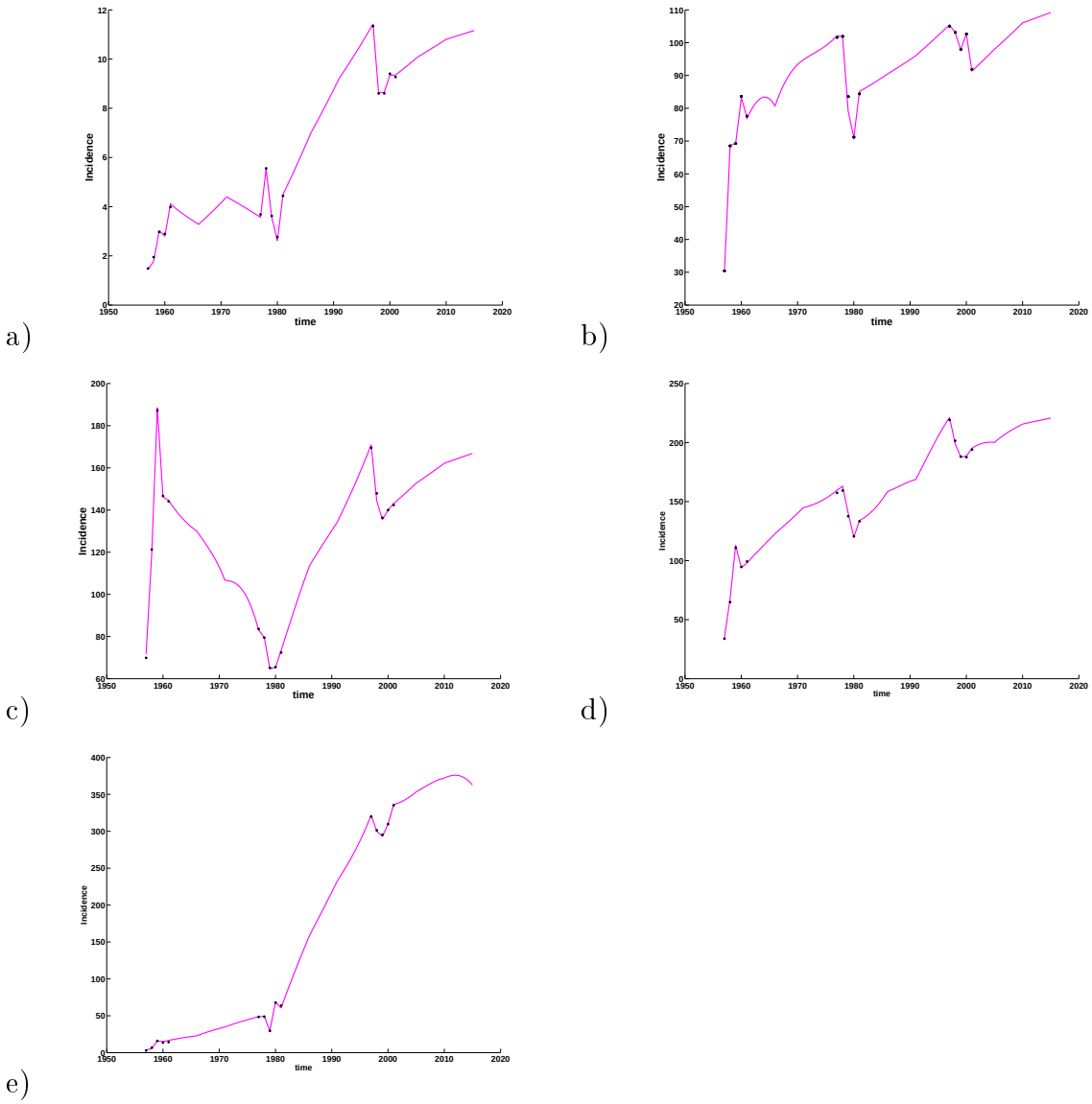


FIG. 4.12. The best model fit to the ages (a) 7 years, (b) 20 years, (c) 35 years, (d) 55 years, (e) and 75 years.

The projected incidence for Taiwan is depicted in FIG. 4.13 and the projection is until 2015. The projection for individuals 65 years and older reflects a sharp increase until 2015, while the other age groups reflect a slight increase in projection with the passage of time.

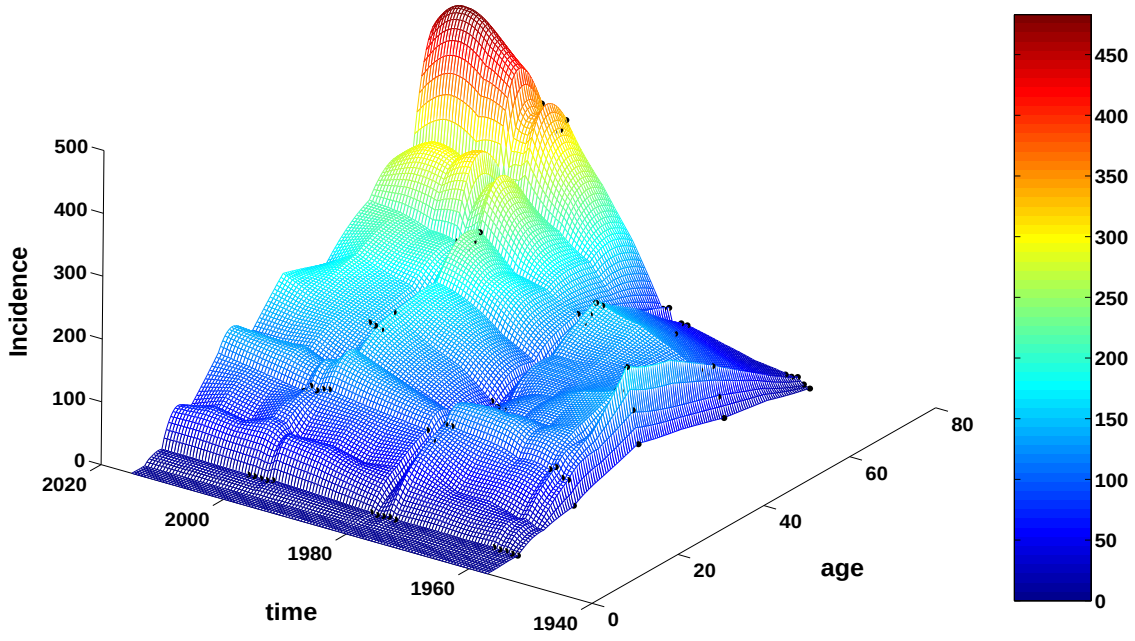


FIG. 4.13. The projection from 1950-2015 of the best fit of the model to the Taiwan data.

FIG. 4.14 depicts the projection for active TB disease incidence in Taiwan as a mesh colour map.

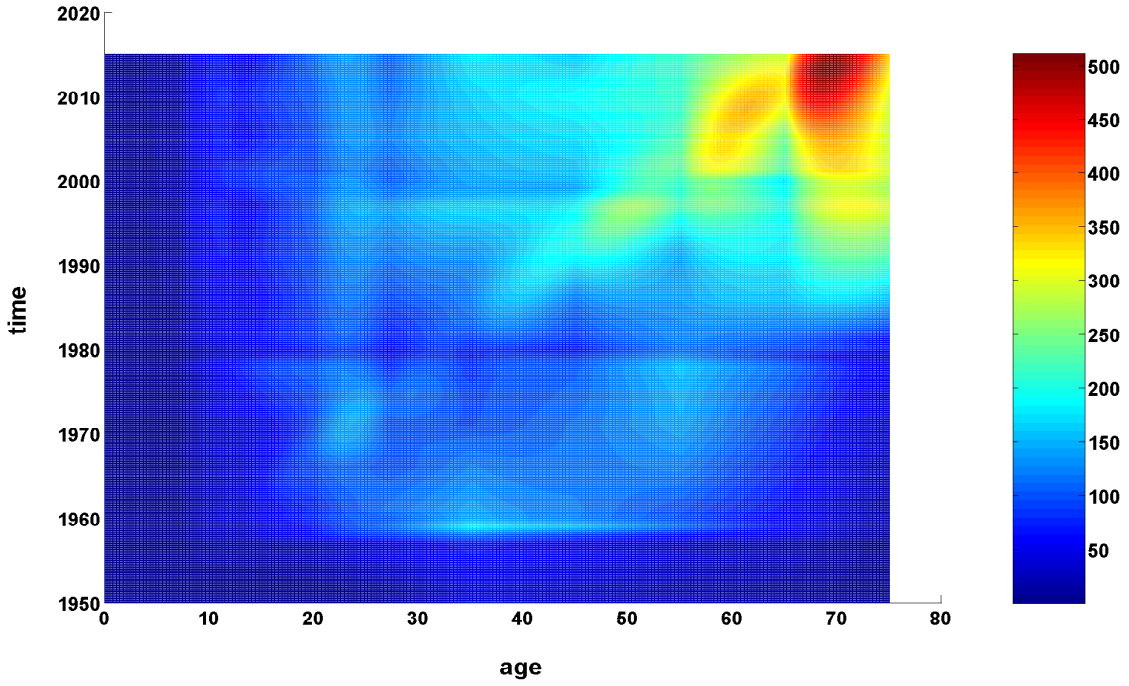


FIG. 4.14. The projection from 1950-2015 of the 5 age groups for active TB disease incidence in Taiwan, considered as a mesh colour map.

4.3 Summary

The TB model developed in Chapter 3 is fitted to the data from Taiwan. The estimation of parameter values and their source is discussed. Assumptions are made to establish the initial and boundary conditions. Appropriate sensitivity analysis was done to investigate the sensitivity of the input parameters. The modifications of the contact rates for each age group are discussed as these were crucial in the model fitting process. Various graphs are produced to show the relationship between incidence, age and time. Projections to determine the future trends of the epidemic are also presented.

In the next chapter we carry out similar simulations, but now on data from Cape Town Metropole.

Chapter 5

Model fit to Cape Town Metropole data

The data set of active TB disease notification rates for Cape Town Metropole was obtained from City of Cape Town Health Department and is displayed in TABLE. 5.1. The data was collected from the TB clinics of the Cape Town Metropole. The Cape Town Metropole includes the regions: Blaauwberg, Cape Town Central, Greater Athlone, Helderberg, Khayelitisha, Mitchell's Plain, Gugulethu, Nyanga, Oostenberg, South Peninsula, Tygerberg Eastern and Tygerberg Western. Data is available for the years 2003-2009 and categorised into the following age classes; 0-4, 5-14, 15-24, 25-34, 35-44, 45-54, 55-64, 65-74 and 75 or older. Data before the year 2003 is deemed to be unreliable.

A distinction in the data is made between pulmonary and extrapulmonary TB. However, we consider only pulmonary TB for our study. The pulmonary TB cases are subject to either a positive smear test, negative smear test or no smear test. The data set is significant since the impact of HIV on this data set can be assumed to be present with a 15% HIV prevalence [1, 25]. This data set is certainly important to our study since the data was collected over time and different age groups.

		Age								
		0-4	**5-14**	**15-24**	**25-34**	**35-44**	**45-54**	**55-64**	**65-74**	**> 74**
Y	**2003**	1812	975	3529	5766	4325	2180	859	257	81
e	**2004**	1791	1004	3771	6367	4637	2183	855	255	100
a	**2005**	1929	1009	4115	6998	5151	2596	939	344	119
r	**2006**	2506	1209	4772	7784	5859	3197	1171	357	135
	2007	2516	1096	4656	7886	6019	3230	1215	373	148
	2008	2830	1120	4564	8455	6589	3543	1299	387	147
	2009	2847	1118	4591	8510	6498	3682	1448	386	147

TABLE. 5.1. Notification rates for active TB disease of Cape Town Metropole from 2003-2009, divided in 9 age intervals.

The data is modified in the same way as that of the Taiwan data apart from the fact that the oldest age interval of the Cape Town Metropole data set is taken to be 10 years. The modified data for Cape Town Metropole is displayed in TABLE. 5.2.

		Age								
		2	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**
Y	**2003**	362.4	97.5	352.9	576.6	432.5	218.0	85.9	25.7	8.1
e	**2004**	358.2	100.4	377.1	636.7	463.7	218.3	85.5	25.5	10
a	**2005**	385.8	100.9	411.5	699.8	515.1	259.6	93.9	34.4	11.9
r	**2006**	501.2	120.9	477.2	778.4	585.9	319.7	117.1	35.7	13.5
	2007	503.2	109.6	465.6	788.6	601.9	323	121.5	37.3	14.8
	2008	566	112	456.4	845.5	658.9	354.3	129.9	38.7	14.7
	2009	569.4	111.8	459.1	851	649.8	368.2	144.8	38.6	14.7

TABLE. 5.2. Modified notification rates for active TB disease of the Cape Town Metropole from 2003-2009.

The assumptions for the boundary condition as a function of time and the initial conditions as a function of age are based on the same reasoning as that of the Taiwan data. We only modified them according to the total population size of Cape Town Metropole and distributed the total population differently in disease classes. We considered the year 1996 as the initial year. The boundary condition for the susceptibles is obtained by taking the population of 1996 in Cape Town Metropole, multiplying it by 0.29 which is the proportion of the individuals aged up to 15 years and then divide it by 15 which is the length of the

age interval to which that population is applicable. The boundary condition is also made an increasing function of time in terms of the average annual growth rate of the population [13].

The boundary condition for the Cape Town Metropole is

$$S(0,t) = 49300\exp[0.029(t-1996)].$$

We assume that the initial prevalence of active TB disease is 10% [27],

while the initial conditions are:

$$\begin{aligned} S(a,0) &= \frac{36628}{1+0.0185a^2}, \\ V(a,0) &= \frac{146512}{1+0.0185a^2}, \\ L(a,0) &= 9508[\exp(0.02a)-1], \\ I(a,0) &= 1.7482a(100-a), \\ R(a,0) &= 2.9752a(100-a). \end{aligned}$$

Similarly to the Taiwan data, the natural mortality rate is approximated by a Gompertz function which is based on Ukraine mortality data investigated by Brooks-Pollock *et al.* [11]. The values of the parameters in Table 5.3 are either assumed or approximated from the literature.

	Age				
Parameter	**At Birth**	**2**	**10**	**20**	**30**
μ	$0.0003\exp(0.0735 \times a)$ [11]				
ψ	0.75	0.675	0.08651251	0	0
p	0.8*	0.85*	0.9*	0.88*	0.9*
r_1	0.009	0.0075	0.0065	0.9	1.8
r_2	0.0135	0.00715	0.002715	0.08	0.41
ρ	0.5**	0.57**	0.65**	0.2875**	0.65**
ϕ	0.04**	0.06**	0.075**	0.035**	0.2**
δ	0.14*	0.07*	0.014*	0.0315125*	0.05125*
σ	5×10^{-9}**	1.5×10^{-8}**	5×10^{-8}**	2×10^{-6}**	0.0001**
θ	0.05	0.25	0.45	0.55	0.65

	Age				
Parameter	**40**	**50**	**60**	**70**	**80**
μ	$0.0003\exp(0.0735 \times a)$ [11]				
ψ	0	0	0	0	0
p	0.92*	0.93*	0.94*	0.95*	0.97*
r_1	2.9	3.25	3.5	3.25	2.9
r_2	0.415	0.42	0.35	0.32	0.3
ρ	0.47**	0.38**	0.39**	0.4**	0.15**
ϕ	0.13**	0.17**	0.21**	0.14**	0.3**
δ	0.065*	0.08*	0.1038*	0.3*	0.4*
σ	0.00022**	0.00025**	0.00028**	0.0004**	0.00045**
θ	0.68	0.72	0.75	0.85	0.95

TABLE. 5.3. Parameter values for different ages. Sources: All the parameter values with * are approximated from Liu *et al.* [28], while the parameter values with ** are approximated from Cohen *et al.* [14]. The natural mortality rate is obtained from Brooks-Pollock *et al.* [11]. The other parameters are based on assumptions.

We fitted the effective contact rate $\beta-$parameter as shown in TABLE. 5.4 to the data by changing them to fit the data.

	Age				
	At Birth	**2**	**10**	**20**	**30**
Maximum	0.751	2.8701	1.795	0.2518	0.2732
Minimum	0.082	0.3122	0.102	0.0364	0.0302
Average	0.5063	2.104	1.2749	0.1821	0.2006

	Age				
	40	**50**	**60**	**70**	**80**
Maximum	0.25	0.1875	0.1238	0.1143	0.6058
Minimum	0.0102	0.0041	0.0014	0.0001	0.00001
Average	0.1237	0.0715	0.0387	0.0375	0.1192

TABLE. 5.4. Transmission rates (β's) for the different age groups.

5.1 Sensitivity Analysis

Similarly, to Chapter 4 Latin Hypercube Sampling and Partial Rank Correlation Coefficients (PRCCs) were used with a 1000 simulations per run. The removed individuals of age 80 years were chosen in our sensitivity analysis as portrayed in FIG. 5.1. This analysis was extended to all the age groups for different parameter values that were age specific. The results were however the same. We thus focus on the analysis presented here as representative of the sensitivity analysis of the model.

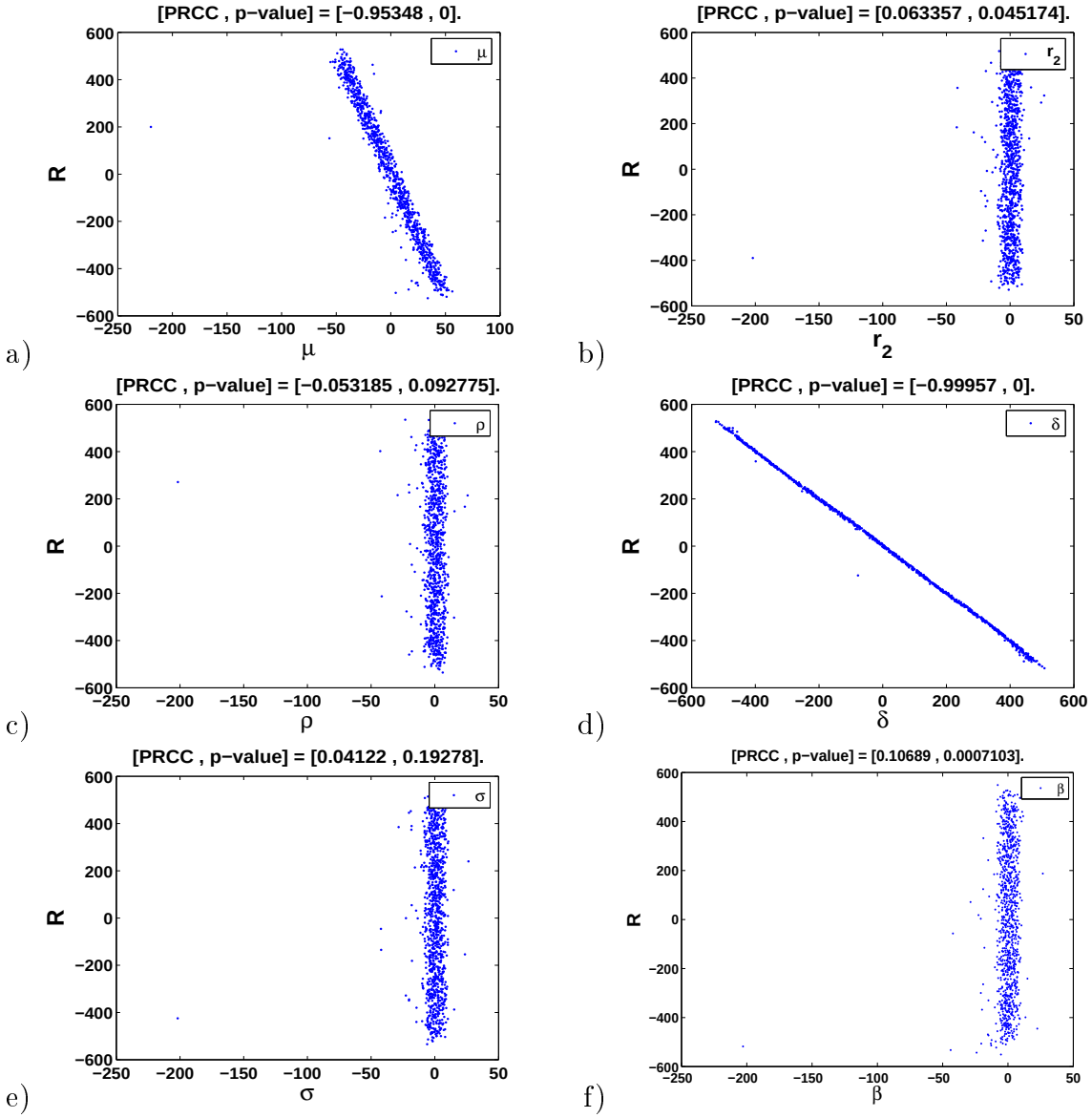


FIG. 5.1. PRCC scatter plots at time 600 years of the most significant parameters which are (a) μ, (b) r_2, (c) ρ, (d) δ, (e) σ and (f) β. The residuals of the linear regression between the rank-transformed values of the output of the removed individuals against rank-transformed values of the significant parameters.

Similarly to Chapter 4, FIG. 5.2 represents the bar plot of the PRCCs related to the different parameters of Cape Town Metropole.

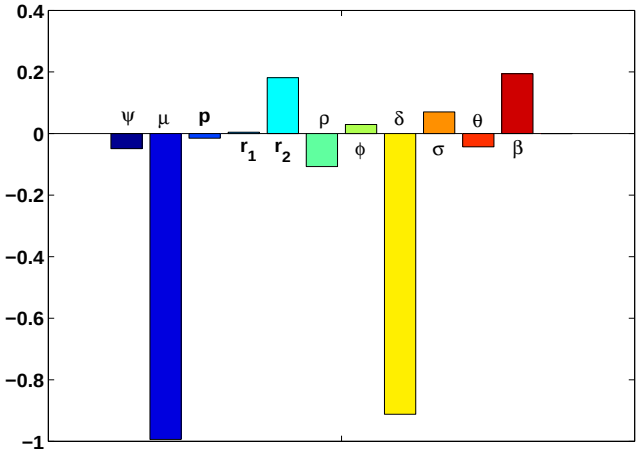


FIG. 5.2. PRCC bar plot of the input parameters with sample size 1000. The parameters with the most significant PRCCs are μ, r_2, ρ, δ, σ and β.

If we take $h = 0.5$ and only assign values for the $\beta's$ of the ages and years corresponding to the data points, with the $\beta-$values being linearly connected to each other, the simulation results do not reflect or fit to the data. Incidence of active TB disease is sensitive to the changes in δ and μ which implies that they could be changed a little and produce different and realistic results. However, the incidence of active TB disease is slightly sensitive to the change in input parameters r_2, ρ, σ and β.

The data of TABLE. 5.2 is plotted in FIG. 5.3. The decreased incidence for age group 5-14 is clearly observable.

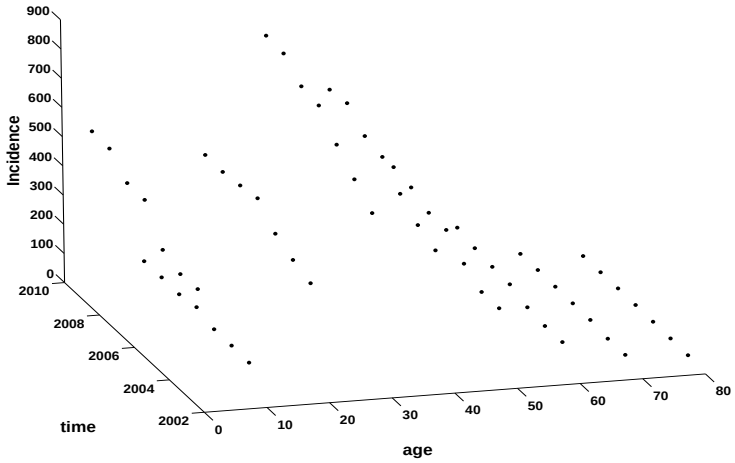


FIG. 5.3. The data of Cape Town Metropole plotted in a three dimensional space with respect to time and age.

The simulations are considered from different view angles and similarly as shown in Figures (5.4) to (5.7) to obtain a sound insight in the age-specific incidence of Cape Town Metropole. Similarly to Taiwan, we now use the discretised model (3.31)-(3.35) and fit it to the data plotted in FIG. 5.3. By an iterative process, that involves changing of parameter values for each age group we obtain the mesh fit depicted in FIG. 5.4. The figure has a corresponding colour bar to show how the incidence changed over years. FIG. 5.4 is specifically a view from the left hand side to observe how incidence of active TB disease appears for the younger ages.

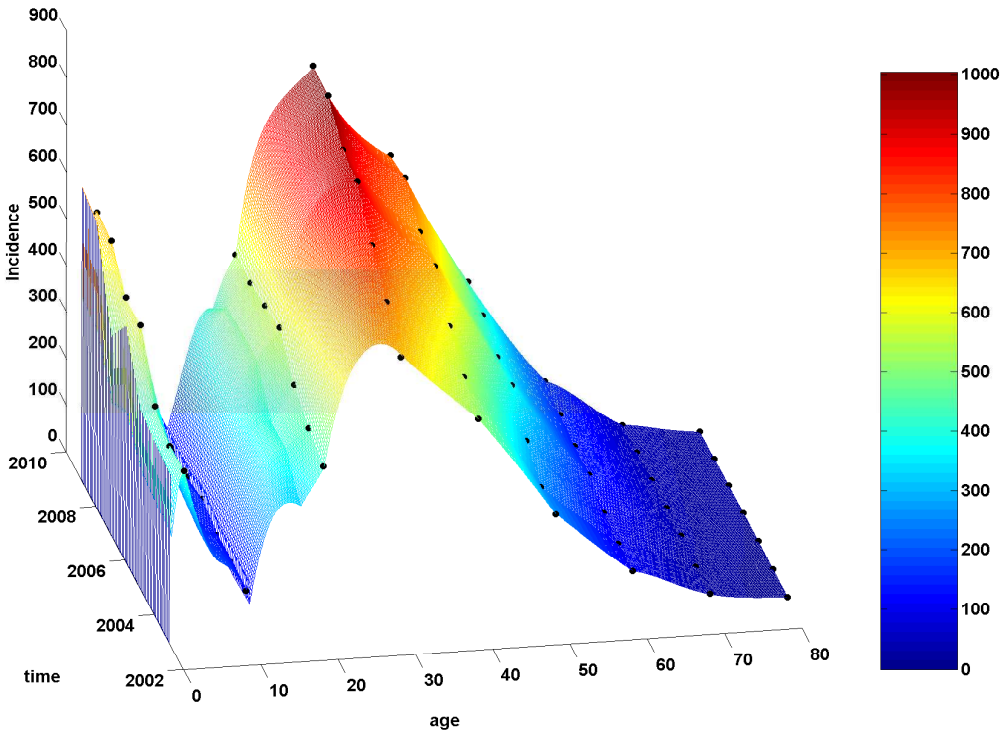


FIG. 5.4. View from the left-hand side of simulation together with the data points of the Cape Town Metropole from 2003-2009 related to the 9 age groups. For the interpretations of colour in the figures the reader is referred to the electronic copy of the thesis.

Although FIG. 5.4 gives us a good idea of the situation in Cape Town Metropole, it may be necessary to look at the same figure from different views. If we rotate the three dimensional FIG. 5.4 in such a way that incidence and age appear on the vertical axes, we can clearly observe how the incidence appears before and after the age of 30 years. FIG. 5.5 is specifically a view from the right hand side to observe how incidence of active TB disease appears for the older ages.

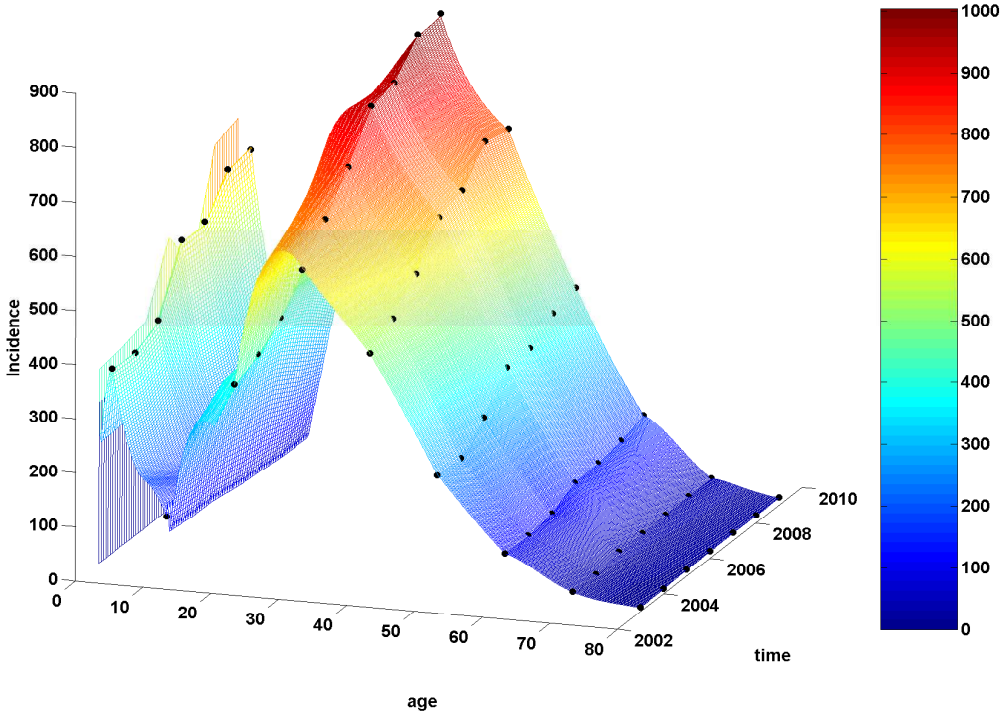


FIG. 5.5. View from the right-hand side of simulation together with the data points of the Cape Town Metropole from 2003-2009 related to the 9 age groups.

A view from the top and the bottom of the three dimensional diagram FIG. 5.4 is shown in figures (5.6) and (5.7). Not many of the data points lie far away from the mesh.

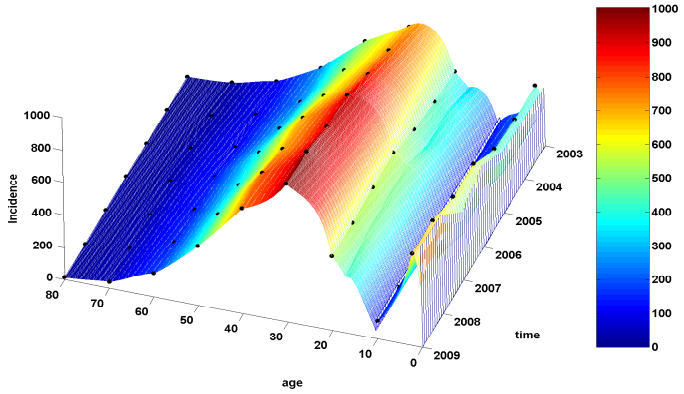


FIG. 5.6. A general view of the simulation together with the data points of the Cape Town Metropole from 2003-2009 related to the 9 age groups.

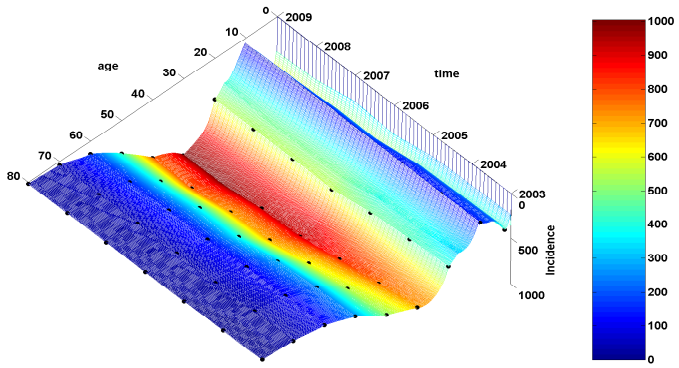


FIG. 5.7. Best fit of the model to the Cape Town Metropole data, which is considered from below.

If we focus on the change in incidence with age, we have the plot FIG. 5.8. It clearly show that TB incidence in Cape Town Metropole is high at the infant stage and then decreases to age 15 years and then increases to the age 35 years after which it decreases again.

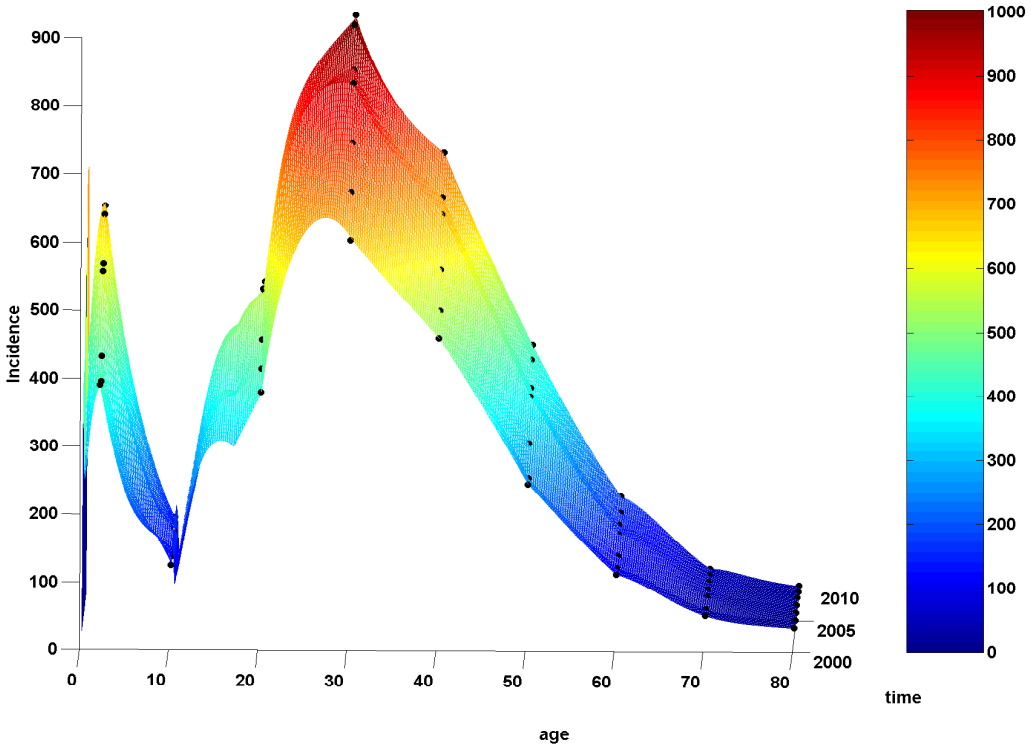


FIG. 5.8. The best fit of the model to the Cape Town Metropole data, which is considered along ages.

The change in incidence over time from 2003 to 2009 is clearly depicted in FIG. 5.9. The figures show that in all age groups TB incidence steadily increased although it is more remarkable for the very young children and the economically productive age groups.

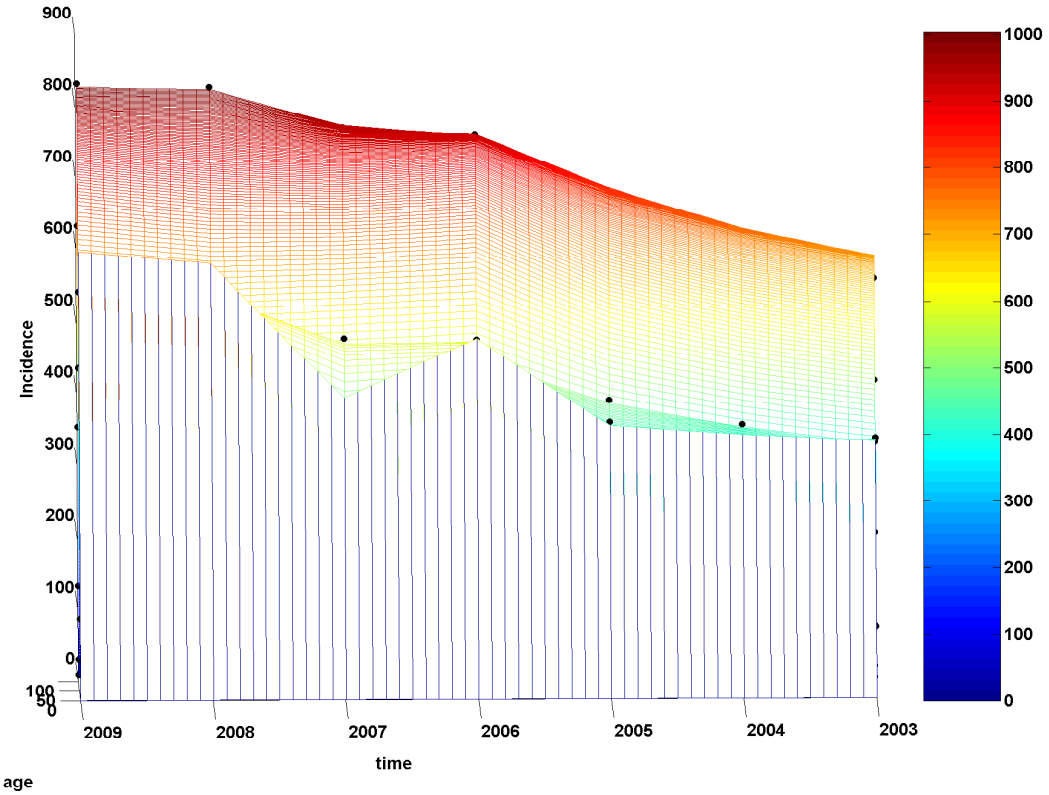


FIG. 5.9. Best fit of the model to the Cape Town Metropole data, which is considered along time in years.

Figures (5.6) to (5.9) can be summarised into a simple and very informative mesh colour map shown in FIG. 5.10. The incidence can be interpreted from the adjacent colour bar. It is clear that the incidence is high for the very young children and the economically productive age groups. It is however low for the old ages.

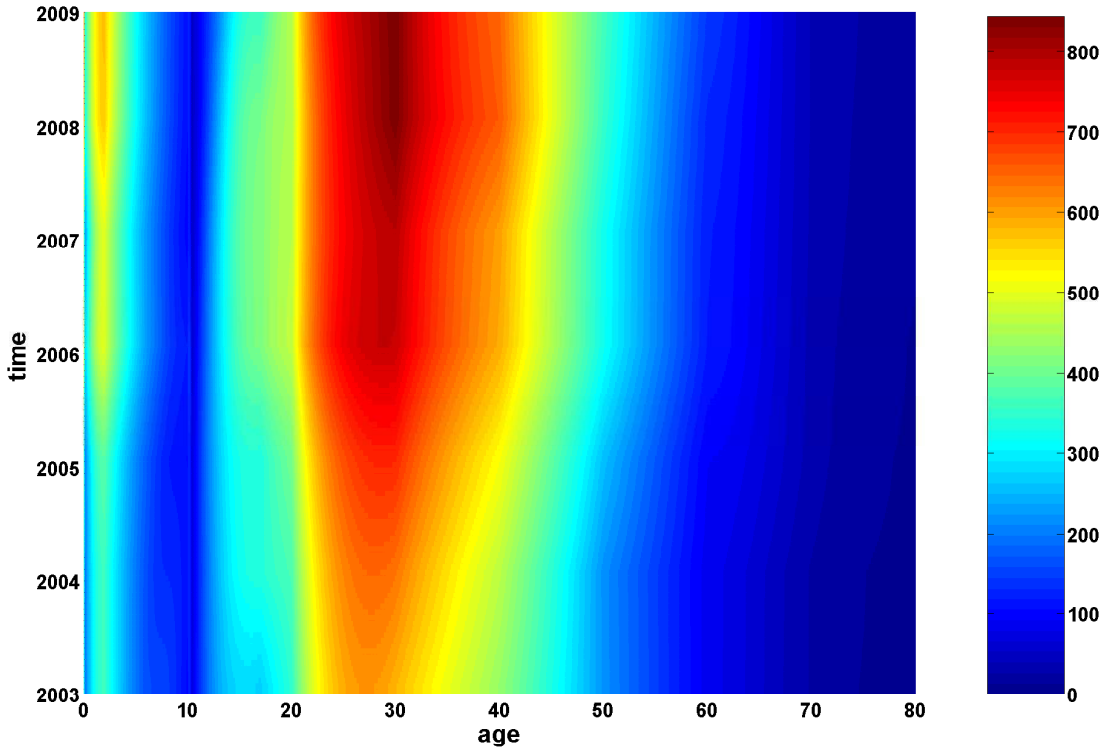
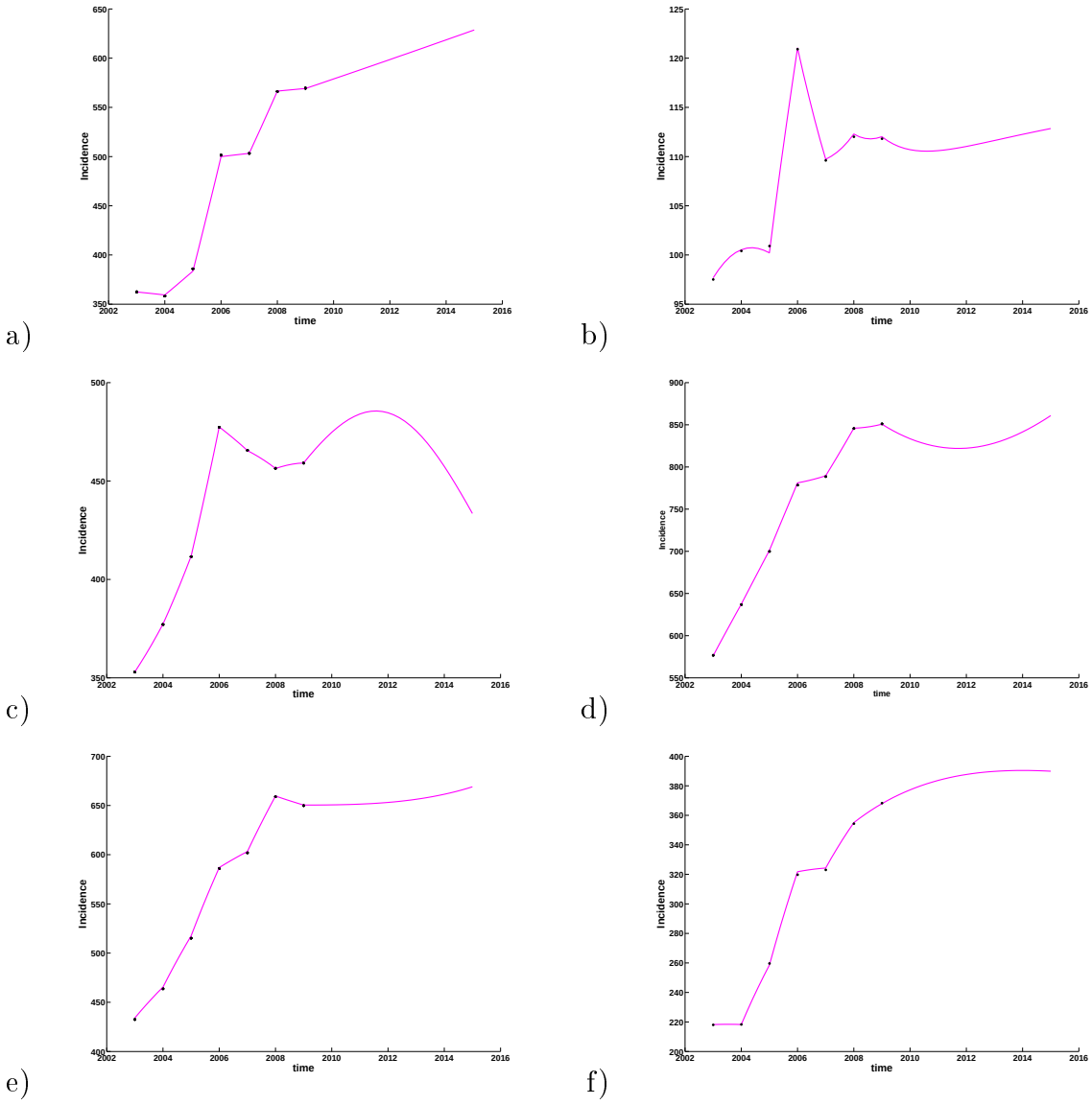
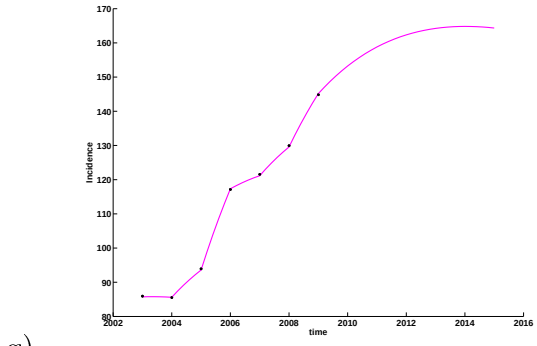


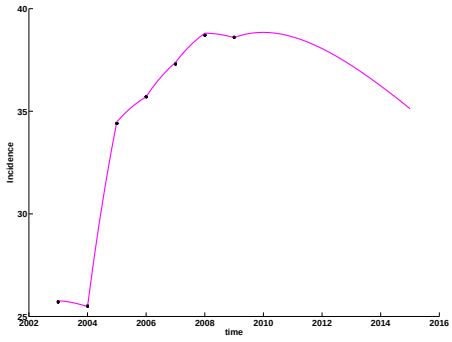
FIG. 5.10. Best fit of the model to the Cape Town Metropole data, considered as a mesh colour map.

The simulations of incidence over time for each age group are illustrated in FIG. 5.11. It shows that all the age groups have had an increasing incidence over time. However, there is a remarkable increase over time in the incidence rates for individuals in the 0-4 year, 25-34 year, 35-44 year and 45-54 year age groups.

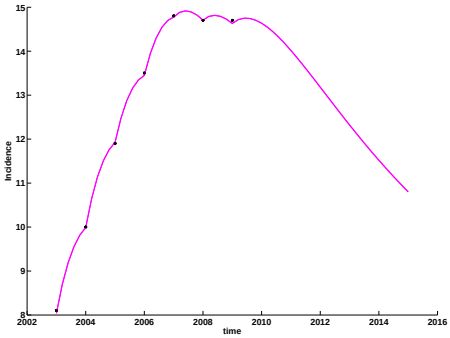




g)



h)



i)

FIG. 5.11. The best model fit to the ages (a) 2 years, (b) 10 years, (c) 20 years, (d) 30 years, (e) 40 years, (f) 50 years, (g) 60 years, (h) 70 years and (i) 80 years.

FIG. 5.12 depicts the projection until 2015 for active TB disease incidence in Cape Town Metropole as a mesh colour map. All the age groups' projections reflect a significant increase in active TB disease incidence. However, the projected incidence rates for individuals in the 0-4 year, 25-34 year, 35-44 year and 45-54 year age groups reflect the most prominent increase in incidence rates over time. It is important to note that the incidence is increasing with decreasing age groups. A comparison of FIG. 5.10 and FIG. 5.12 shows that the highest incidence of TB occurs amongst the 30 year olds in FIG. 5.10 and is projected to be the highest amongst the 25 olds in 2015.

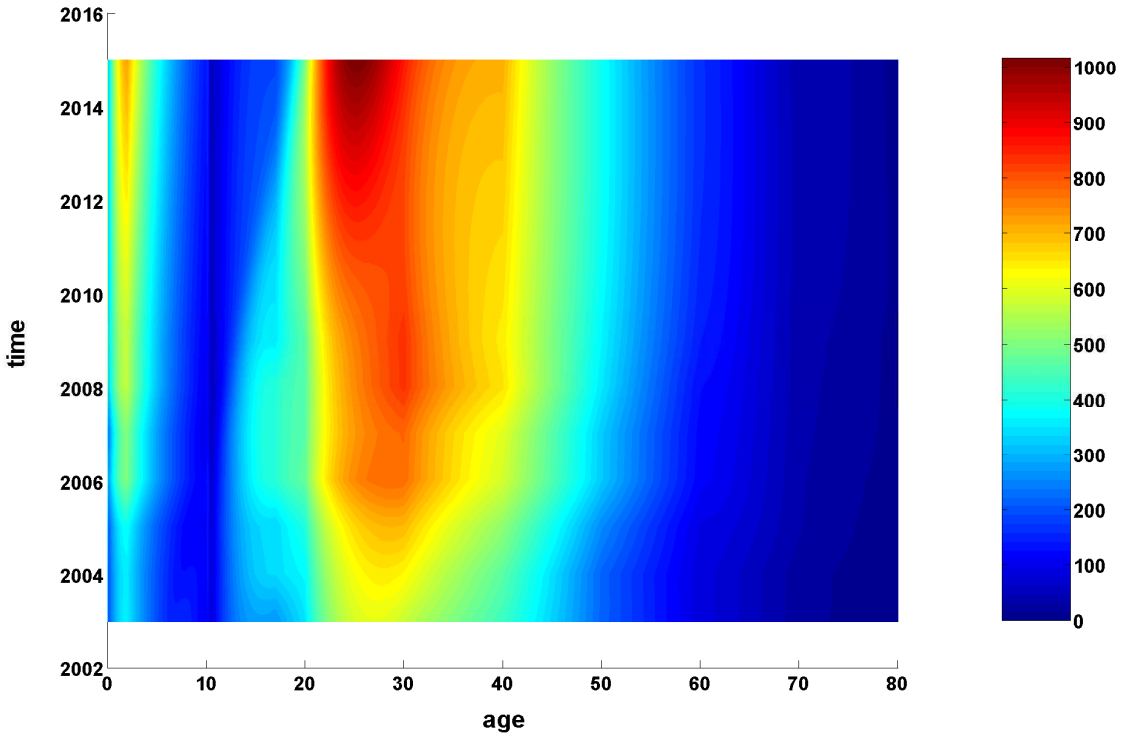


FIG. 5.12. The projection for active TB disease incidence in Cape Town Metropole, considered as a mesh colour map.

5.2 Summary

Similarly to Chapter 4, the TB model developed in Chapter 3 is fitted to the data from Cape Town Metropole. We took the same approach as in Chapter 4 regarding the estimation of parameter values, initial and boundary conditions and sensitivity analysis. Likewise, the same emphasis was laid on the importance of effective contact rates. The relationship between incidence, age and time and its projections are also viewed pictorially to obtain better insight of the disease.

Chapter 6

Discussion

Clinical data indicates that the HIV epidemic has played a role in shifting the TB epidemic to younger age groups. In 1996-1997, the largest number of TB notifications was in the age group 40-49. By 2004 it had shifted to the 25-30-year-olds [36]. This thesis was built on the premises that mathematical modeling could improve the understanding of TB dynamics in an area with high HIV prevalence such as the Western Cape province of South Africa. Using a data from the City of Cape Town, Health Department, the application of a simple but reliable TB model is dealt with.

In Chapters 4 and 5 we fitted our mathematical model to the Taiwan and Cape Town Metropole data, respectively. We compare two data sets of active TB disease incidence rates in our study which are the Western Cape Tuberculosis Programme from the City of Cape Town Health Department and the data set from Yu *et al.* [51]. The comparison of the two data sets allows us to determine and quantify the effects of HIV has on TB dynamics. This is particular of importance as HIV has been labeled as the main driver of TB in sub-Saharan Africa.

6.1 Comparison of Taiwan and Cape Town Metropole Notification Rates

It is important that the age intervals of Taiwan and Cape Town Metropole correspond with each other in order to make comparison possible. Since we work with incidence rates for a

whole population it is important to take the population size into account. The distribution of the age groups are multiplied by the population at the time in order to obtain the population for each age group [13, 40]. If the distribution of the data does not correspond to that of the study, we assume a uniform distribution of individuals within age intervals. The age intervals are then merged and split depending on how intervals appear in the literature.

The age distributions of Taiwan for 1965, 1980 and 1995 are obtained and we assume it applies to the respective intervals of 1957-1961, 1977-1981 and 1997-2001 [15]. The age distribution of 2005 for South Africa is assumed to hold for the Cape Town Metropole from 2003-2009 [13]. The incidence rates for the corresponding age groups are divided by the total population for each age group in number of thousands to obtain a number of individuals per 1000 who are subject to the incidence of active TB disease. The average of these incidence rates over the time intervals 1957-1961, 1977-1981 and 1997-2001 of Taiwan and 2003-2009 of Cape Town Metropole for each age group is obtained. The values of Cape Town Metropole are divided by the 3 sets of values of Taiwan to get the comparative values in terms of Cape Town Metropole.

After we take the population size of both regions in consideration, the incidence of active TB disease for the economically productive age groups of Cape Town Metropole is considerably higher than that of Taiwan. It is very clear from TABLE. 6.1 that HIV has a significant effect on the active TB disease incidence rates of very young children and the economically productive individuals.

Year interval	**0-14**	**15-24**	**25-44**	**45-64**	**$\geq$ 65**
1957-1961	337.365	18.159	13.622	5.848	9.36
1977-1981	289.8	24.892	45.495	7.323	2.31
1997-2001	115.172	23.096	35.193	7.614	0.812

TABLE. 6.1. The Cape Town Metropole notification rates are compared with different time intervals of Taiwan related to the corresponding age intervals. The Taiwan data are used as baseline. The values show how much higher the incidence of active TB disease is for Cape Town Metropole than that of different time intervals of Taiwan

6.2 Comparison of the Taiwan and Cape Town Metropole Parameter Values

Bacaër *et al.* [4] differentiates between physiological and social parameters of TB mathematical models. The physiological parameters are parameters which are considered to be the same across the world and in our model specifically they are $\Pi(t)$, $\mu(a)$, $\theta(a)$, $p(a)$, $r_1(a)$, $r_2(a)$ and $\sigma(a)$. On the other hand social parameters are directly associated with living conditions and HIV and in this model they are $\psi(a)$, $\rho(a)$, $\phi(a)$, $\delta(a)$ and $\beta(a,t)$.

The effective contact rates of Cape Town Metropole are on average higher than those of Taiwan. The other parameters were adjusted according to literature related to TB for regions at times similar to situations for these two regions at their respective time frames.

The parameter values of $\psi(a)$, $p(a)$ and $\theta(a)$ are of about similar size for the corresponding ages when comparing the Taiwan and Cape Town Metropole cases, while for $r_1(a)$ and $\phi(a)$ their values applicable to Cape Town is twice as high than that of Taiwan. The respective parameters $r_2(a)$, $\rho(a)$, $\delta(a)$ and $\sigma(a)$ are 3, 0.65, 20 and 0.8 times higher for Cape Town Metropole than that for Taiwan.

According to the simulations the average β-values across the times are within the range of [14.997, 1581.886] for the ages up to 15 years. The probable reason for this anomaly is that the first interval of the Taiwan data set is applicable to the first 15 years, while for the Cape Town Metropole data this similar age interval is stratified into two age intervals. In the case of Taiwan, no distinction is made between the infant years and the "golden ages" which implies that the "golden ages" offset the infant ages notification rates of active TB disease.

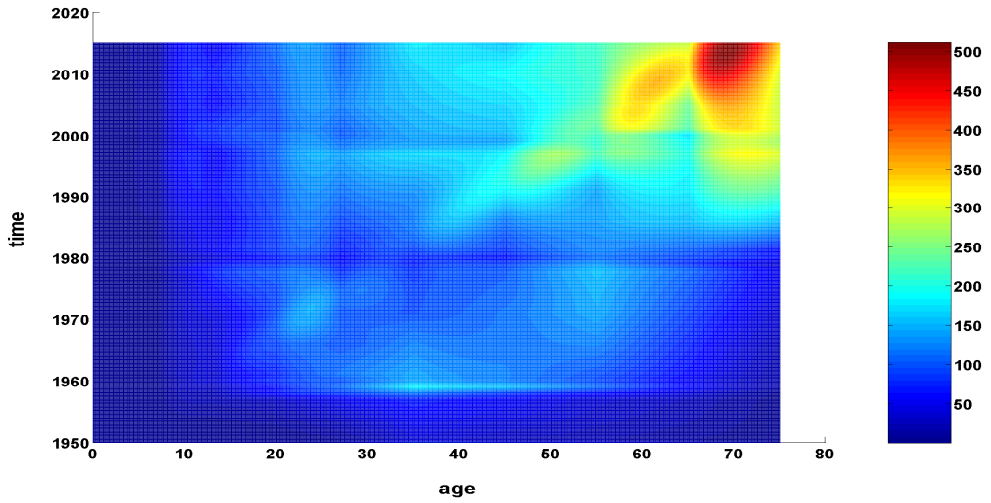
According to TABLE. 6.2, the β-values for Cape Town is considerably higher than that for Taiwan up to the age 50 years. After 50 years of age the β-values are very comparable in size. The reason for that is the high HIV prevalence among the economically productive population of Cape Town Metropole. The difference in age-specific active TB disease incidence over time for Taiwan and Cape Town Metropole is depicted by FIG. 6.1.

Age Interval	Range of comparison size
$[15, 27]$	$[16.987,\ 22.959]$
$[27, 40]$	$[15.496,\ 22.959]$
$[40, 50]$	$[6.245,\ 15.496]$
$[50, 60]$	$[2.12,\ 6.245]$
$[60, 75]$	$[1.5207,\ 2.12]$

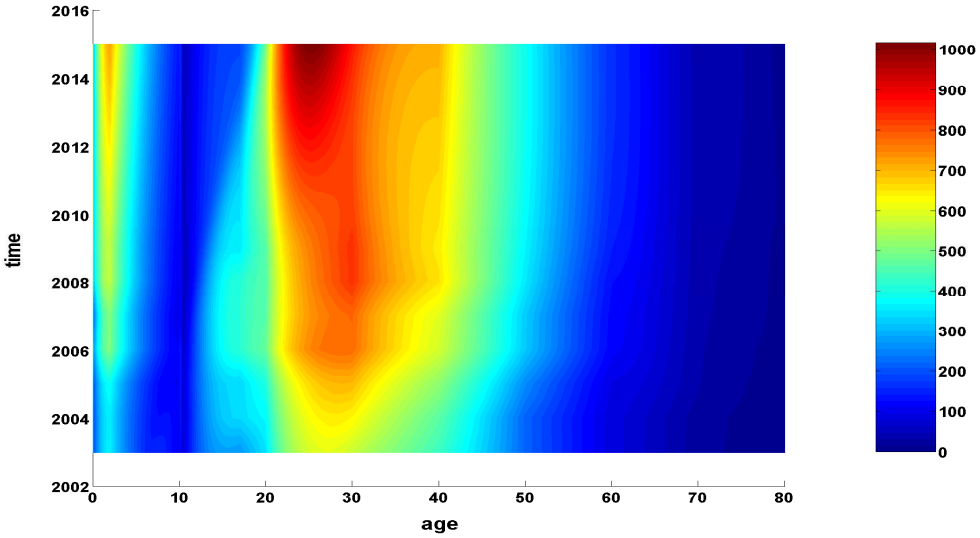
TABLE. 6.2. The comparison of the Cape Town Metropole and Taiwan β-values corresponding to age intervals.

6.3 Comparison of the Taiwan and Cape Town Metropole Fits

At first sight, the age distribution of the notification rates for active TB of Cape Town Metropole seems to be similar to that for Taiwan for the period 1957-1961. After those years the dynamics for the two regions differ increasingly with passage of time. We give the critical and most important results of this study that shows that TB incidence is high in the economically active population in the Cape Town Metropole when compared to Taiwan, where the incidence is high amongst the elderly. Figure 6.1 shows a comparison of the projected incidence over time for the two regions.



a)



b)

FIG. 6.1. The projections until 2015 of active TB disease incidence for (a) Taiwan and (b) Cape Town Metropole.

The implication is that the age groups that have the highest incidence rates of active TB disease should make the largest contribution to the transmission of TB. In the Cape Town Metropole, higher incidence rates of TB disease are as a result of HIV infection. The sexually active population which are individuals between the ages of 15 and 49 years are affected by HIV infection. During the fitting process, it is interesting to note that there

was a direct relationship between mortality due to TB disease and the infection rates. We noted that the higher the transmission of TB, the higher the disease induced deaths. This is reasonable especially when we consider the assumption on which the model is built. A comparison of FIG. 5.10 and FIG. 5.12 shows that TB incidence will increase in the lower age groups of the population.

6.4 Limitations

It is important to take epidemiological modeling from theoretical realm to applications to real-world scenarios. While this does not invalidate the usefulness of mathematical models as tools to understand dynamical systems, it is clearly evident that even very simple models based applied to limited data can actually give valuable insights. Like every modelling exercise, this work is not without limitations. It is important to note the comparisons, especially the data from Taiwan and data from the Metropole, raises a considerable number of questions. Firstly, the data from Taiwan is on a national scale while our data is at a city level and its surroundings. This means that the sizes of the populations under consideration differ significantly. Secondly, TB is a disease that is usually influenced by the socio-economic conditions of individuals. There are differences, economically and socially between the two areas. Thirdly, the most important aspect to consider is HIV/AIDS. Taiwan has had no recorded HIV infection during the periods in which the data has been collected. This disparity is what we used in this work to try and determine the role of HIV infection on TB incidence. We also assumed that the differences in living conditions is negligibly small on the effect of TB. However, the difference in living conditions should also be further considered but although it is outside the scope of this dissertation.
Despite these limitations, some interesting results were obtained. There is however room to improve this work. The study could be extended to the whole of South Africa and comparisons with Taiwan would become more reasonable. This will remain a work in progress for as long as these epidemics endure.

Appendix A

Matlab Code and Data for Simulations

A.1 Simulations for Taiwan Data

```
clear all

h = 0.5;   % h is the grid size
begin_year = 1950;
age = 0:h:75;
time = begin_year:h:2015;
[X,Y] = meshgrid(age,time);

% Variable which runs the code for a number of times to provide convergence
numruns = 100;

% Declare all the classes and incidence by means of zero matrices
S = zeros(length(age),length(time));
V = zeros(length(age),length(time));
L = zeros(length(age),length(time));
I = zeros(length(age),length(time));
R = zeros(length(age),length(time));
Incidence = zeros(length(age),length(time));

% Proportion of Susceptibles vaccinated initially
frac = 0.46;

% Boundary Condition
S(1,:) = 247093*exp(0.0145*(time-1950));

% Initial Conditions
S(:,1) = (1-frac)*457848*(1./(1 + 0.0185*age.^2));
V(:,1) = (frac)*S(:,1);
L(:,1) = 19017*(exp(0.02*age)-1));
I(:,1) = 2.9058*(-age + 100).*(age);
```

```
R(:,1) = 5.4123*(-age + 100).*(age);

%Code to find right population size and proportion of classes
sum(S(:,1) + V(:,1)+ L(:,1)+ I(:,1)+ R(:,1))
[sum(S(:,1)) sum(V(:,1)) sum(L(:,1)) sum(I(:,1)) sum(R(:,1))]
   /sum(S(:,1) + V(:,1)+ L(:,1)+ I(:,1)+ R(:,1))

mu = 0.0003*exp(0.0735*age);   % natural mortality rate

% Ages linked to the corresponding entries of parameter vectors
ages = [0 7 20 35 55 75];

% Ages  0   7   20   35   55   75

% Parameters vectors corresponding to ages
psi = [0.75 0.08651251235750 0 0.00000000000000000001 0 0.0000000000000000001];
p = [0.8 0.9 0.88 0.92 0.93 0.95];
r1 = [0.0045 0.00325 0.45 1.45 1.75 1.45];
r2 = [0.0045 0.0009051 0.0265 0.1365 0.14 0.1];
rho = [0.75 0.975 0.875 0.9675 0.575 0.6];
phi = [0.02 0.03875 0.0175 0.065 0.1065 0.07065];
delta = [0.007 0.0007 0.0015125 0.0025125 0.005138 0.015];
sigma = [0.000000005 0.00000005 0.000002 0.00022 0.00028 0.0004];
theta = [0.05 0.45 0.55 0.65 0.75 0.95];

% Declare all the transmission matrix by means of a zero matrix
betamatty = zeros(length(time),length(age));
years = [1950 1957:1:1961 1966 1971 1977:1:1981 1986 1991 1997:1:2001 2005:5:2015];

% Specify transmission vectors which entries correspond to:
% Ages   0   3   7   14   20   27   35   45   55   65   75

beta_1950 = [0.000318656875 0.00000019366006 0.00000156875 0.0000094375 0.0001588253125
             0.0015 0.0054693825 0.003125 0.0019007938125 0.0009375 0.00008766194375];
beta_1957 = [0.00020394 0.00390092394 0.011 0.0074484 0.0114981648 0.0111458
             0.011004 0.0086 0.0051009216508 0.004 0.0009610364];
beta_1958 = [0.0002039404 0.0042852 0.013004 0.007012 0.0234 0.01851 0.018
             0.01400004 0.0101312608 0.007008152 0.00141184561036];
beta_1959 = [0.0002860394 0.0048212394 0.02 0.00814168 0.0212868 0.021881272
             0.028 0.0253004 0.01738010866 0.0114 0.004161038];
beta_1960 = [0.00006039404 0.0052394084 0.019004 0.009424 0.0228732 0.0183368
             0.0213708 0.0194866 0.01484 0.010204 0.00418856];
beta_1961 = [0.00005284475 0.0064084475 0.026285 0.00809050935 0.0190155
             0.0200375 0.02065 0.01976499 0.015784563305 0.01197175 0.0049090685];

% Ages   0   3   7   14   20   27   35   45   55   65   75

beta_1966 = [0.000690209025 0.003602109775 0.02101540375 0.012879625 0.01339525
             0.0109627375 0.015147374 0.01952856675 0.0216262068 0.01705 0.00863571225];
beta_1971 = [0.00082850625 0.004600350625 0.0249315625 0.0126832875 0.014220350625
             0.0063375 0.00853125 0.01783309125 0.02682875 0.023075 0.015846025];
```

```
beta_1977 = [0.0007549255 0.001065492705 0.0192 0.0201155 0.01525 0.00435755
             0.00395 0.01291825 0.02907595 0.030286056 0.02645];
beta_1978 = [0.00006254928 0.0003654925 0.027415 0.0204361755 0.015125 0.0040525
             0.003525 0.0118 0.02925 0.032055 0.02737];
beta_1979 = [0.0000469119375 0.00026366195662 0.018646875 0.016320375 0.0116717896875
             0.00367921875 0.002701125 0.009556875 0.02467875 0.0240493125 0.017];
beta_1980 = [0.0000312746275 0.0001775774625 0.0143 0.01013 0.010375 0.00331 0.0025875
             0.006625 0.0205813 0.027469135875 0.03977435];
beta_1981 = [0.0000281471625 0.002947169475 0.02188125 0.0145125 0.01227375
             0.0039423465 0.0027675 0.00792 0.022275 0.0307125 0.0361125];

% Ages   0   3   7   14   20   27   35   45   55   65   75

beta_1986 = [0.0012549255 0.00154925 0.03 0.01155 0.012175 0.0049052259277
             0.00375655 0.00672565 0.02063505 0.04771075 0.10177012955];
beta_1991 = [0.00125492675 0.0015492955 0.0358 0.0135 0.012 0.005648705
             0.00428075 0.0065 0.01521251 0.05125 0.1562];
beta_1997 = [0.00000042549265 0.00005154925 0.04 0.010916255 0.01205
             0.00625 0.00515 0.006336174 0.013 0.045535 0.2055];
beta_1998 = [0.00000035625493 0.0000015492605 0.0305 0.0125762755 0.011625
             0.005148214175 0.00429855 0.00563825 0.011065 0.040625 0.1875];
beta_1999 = [0.0000002549275 0.00000092565492 0.03005 0.01526255 0.0108775
             0.004905 0.0039875 0.0048575 0.0099115 0.0375 0.178];
beta_2000 = [0.0000001254926 0.00000086549271 0.03175 0.0155 0.011245 0.0043775
             0.0040635 0.0052500993375 0.0095 0.031723 0.182];
beta_2001 = [0.00000006254929 0.00000085654925 0.031286 0.015826255 0.009875
             0.0043825 0.0041 0.00551291205 0.00945 0.038406725 0.19];

% Ages   0   3   7   14   20   27   35   45   55   65   75

beta_2005 = [0.0000025492865 0.00000365492805 0.031598615 0.01092655 0.01
             0.0048125 0.0041295506 0.0056376705 0.00862697 0.027336 0.1623];
beta_2010 = [0.0000012549285 0.000003654925 0.031286 0.0118226 0.01005 0.0041875
             0.004075 0.0058625 0.0086175 0.0229 0.1225];
beta_2015 = [0.0000012549288 0.00000365492605 0.03 0.007855 0.0096 0.00380625
             0.0039 0.00455825 0.00837615 0.01756705 0.086562675];

% Assigning the parameter vectors to their respective matrices

bmatty = vertcat(beta_1950,beta_1957,beta_1958,beta_1959,beta_1960,beta_1961,
         beta_1966,beta_1971,beta_1977,beta_1978,beta_1979,beta_1980,beta_1981,
         beta_1986,beta_1991,beta_1997,beta_1998,beta_1999,beta_2000,beta_2001,
         beta_2005,beta_2010,beta_2015);

paramatty = vertcat(psi, p, r1, r2, rho, phi, delta, sigma, theta);

% Code for interpolating parameters linearly between different ages
size_paramatty = size(paramatty);
nrows_paramatty = size_paramatty(1);
paramdiffy = zeros(nrows_paramatty,length(ages)-1);
param = zeros(nrows_paramatty,length(age));
```

```
for i = 1:1:nrows_paramatty
    for j = 1:1:(length(ages)-1)
       paramdiffy(i,j) = ((paramatty(i,j+1) - paramatty(i,j))*h)/
       (ages(j+1) - ages(j));
    end
end

for k = 1:1:(nrows_paramatty)
    vect1 = [];
    for m = 1:1:(length(ages)-2)
        vect1 = horzcat(vect1,paramatty(k,m):paramdiffy(k,m):
        (paramatty(k,m+1)-paramdiffy(k,m)));
    end
    vect1 = horzcat(vect1,paramatty(k,(length(ages)-1)):
    paramdiffy(k,(length(ages)-1)):paramatty(k,length(ages)));
    param(k,:) = vect1;
end

psi = param(1,:);
p = param(2,:);
r1 = param(3,:);
r2 = param(4,:);
rho = param(5,:);
phi = param(6,:);
delta = param(7,:);
sigma = param(8,:);
theta = param(9,:);


ages = [0 3 7 13 20 27 35 45 55 65 75];
bdiffy = zeros(length(years),length(ages)-1);

for i = 1:1:length(years)
    for j = 1:1:(length(ages)-1)
    bdiffy(i,j) = ((bmatty(i,j+1) - bmatty(i,j))*h)/(ages(j+1) - ages(j));
    end
end

% Variable for converting a year number into a year since the initial year
yearsad = years - begin_year;

% Code for interpolating transmission parameters linearly between different ages and times
for k = 1:1:length(years)
    vect1 = [];
    for m = 1:1:(length(ages)-2)
        vect1 = horzcat(vect1,bmatty(k,m):bdiffy(k,m):
        (bmatty(k,m+1)-bdiffy(k,m)));
    end
    vect1 = horzcat(vect1,bmatty(k,(length(ages)-1)):
    bdiffy(k,(length(ages)-1)):bmatty(k,length(ages)));
```

```
    betamatty(((yearsad(k)/h) + 1),:) = vect1;
end

bmdiffy = zeros(length(years)-1,length(age));

for k = 1:1:length(age)
    for n = 1:1:(length(years)-1)
        bmdiffy(n,k) = ((betamatty(yearsad(n+1)/h + 1,k) -
        betamatty(yearsad(n)/h + 1,k))*h)/(yearsad(n+1)-yearsad(n));
    end
end

for k = 1:1:length(age)
    vect = [];
    for m = 1:1:(length(yearsad)-2)
        vect = horzcat(vect,betamatty(yearsad(m)/h + 1,k):bmdiffy(m,k):
        betamatty(yearsad(m+1)/h + 1,k)-bmdiffy(m,k));
    end
    vect = horzcat(vect,betamatty(yearsad(length(yearsad)-1)/h + 1,k):
    bmdiffy(length(yearsad)-1,k):betamatty(yearsad(length(yearsad))/h + 1,k))';
    betamatty(:,k) = vect';
end

betamatty = betamatty';

%Double for loop of the Infectives Class'es first guess.
for j = 2:1:length(time)
    for i = 2:1:length(age)
    lambda = betamatty(i-1,j)*I(i-1,j);
    S(i,j) = (S(i,j-1) + S(i-1,j))/(2+h*(lambda + mu(i-1)+psi(i-1)));
    V(i,j) = (h*psi(i-1)*S(i,j) + V(i,j-1) + V(i-1,j))/
    (2+h*(theta(i-1)*lambda + mu(i-1)));
    R(i,j) = (h*rho(i-1)*I(i-1,j) + R(i,j-1) + R(i-1,j))/
    (2+h*(r2(i-1)*lambda + phi(i-1) + mu(i-1)));
    L(i,j) = (h*lambda*(p(i-1)*S(i,j) + theta(i-1)*V(i,j)+ r2(i-1)*R(i,j)) +
    L(i,j-1)+ L(i-1,j))/(2+h*(r1(i-1)*lambda + sigma(i-1) + mu(i-1)));
    I(i,j) = (h*(lambda*(1-p(i-1))*S(i,j) + (lambda*r1(i-1) + sigma(i-1))*
    L(i,j)+ phi(i-1)*R(i,j)) + I(i,j-1)+ I(i-1,j))/
    (2+h*(delta(i-1) + rho(i-1) + mu(i-1)));
    end
end

% Triple for loop of the convergence for the system of difference equations.
for runs = 1:1:numruns
    for j = 2:1:length(time)
        for i = 2:1:length(age)
            lambda = betamatty(i,j)*I(i,j);
            S(i,j) = (S(i,j-1) + S(i-1,j))/(2+h*(lambda + mu(i)+psi(i)));
            V(i,j) = (h*psi(i)*S(i,j) + V(i,j-1) + V(i-1,j))/
            (2+h*(theta(i)*lambda + mu(i)));
            R(i,j) = (h*rho(i)*I(i,j) + R(i,j-1) + R(i-1,j))/
```

```
            (2+h*(r2(i)*lambda + phi(i) + mu(i)));
            L(i,j) = (h*lambda*(p(i)*S(i,j) + theta(i)*V(i,j)+ r2(i)*R(i,j)) +
            L(i,j-1)+ L(i-1,j))/(2+h*(r1(i)*lambda + sigma(i) + mu(i)));
            I(i,j) = (h*(lambda*(1-p(i))*S(i,j) + (lambda*r1(i) + sigma(i))*
            L(i,j)+ phi(i)*R(i,j)) + I(i,j-1)+ I(i-1,j))/
            (2+h*(delta(i) + rho(i) + mu(i)));
        end
    end
end

% Double for loop to fill out the matrix of incidence.
for j = 2:1:length(time)
    for i = 2:1:length(age)
        Incidence(i,j) = ((1-p(i))*betamatty(i,j)*I(i,j)*S(i,j) + sigma(i)*L(i,j) +
        r1(i)*betamatty(i,j)*I(i,j)*L(i,j) + phi(i)*R(i,j))/(S(i,j) + L(i,j) + R(i,j));
    end
end

% Years and ages related to the Taiwan Data
years = [1957:1:1961 1977:1:1981 1997:1:2001];
ages = [7 20 35 55 75];
U = zeros(length(years),length(ages));

% Taiwan Data
U(1,1:5) = [22 300 1380 668 57];
U(2,1:5) = [29 681 2411 1289 133];
U(3,1:5) = [44 684 3699 2191 313];
U(4,1:5) = [43 829 2909 1877 268];
U(5,1:5) = [59 766 2846 1961 282];
U(6,1:5) = [55 1013 1666 3138 963];
U(7,1:5) = [83 1015 1583 3174 971];
U(8,1:5) = [54 831 1296 2736 586];
U(9,1:5) = [41 704 1295 2384 1342];
U(10,1:5) = [66 837 1436 2645 1264];
U(11,1:5) = [170 1050 3389 4382 6395];
U(12,1:5) = [129 1031 2957 4029 6023];
U(13,1:5) = [129 979 2726 3759 5903];
U(14,1:5) = [141 1026 2799 3752 6192];
U(15,1:5) = [139 918 2847 3880 6702];

% Incorporate the "Unknown" cases in the data set according to ages
Unknown = [30 26 81 49 73 20 25 28 63 49 0 0 0 0 0];
for i = 1:1:length(years)
    for j = 1:1:length(ages)
        U(i,j) = U(i,j)*(1 + (Unknown(i)/sum(U(i,1:5))));
    end
end

% Length of age intervals
int_length = [15 10 20 20 20];
```

```
% Divide each data entry by its corresponding age interval
for i = 1:1:length(years)
    U(i,1:5) = U(i,1:5)./int_length;
end

% Code for scatterplot
yearsmat = zeros(length(years),length(ages));
for i = 1:1:length(years)
    yearsmat(i,:) = ones(1,length(ages))*years(i);
end

Uvec = [];
agesvec = [];
yearsvec = [];
for i = 1:1:length(years)
    Uvec = horzcat(Uvec,U(i,:));
    agesvec = horzcat(agesvec,ages);
    yearsvec = horzcat(yearsvec,yearsmat(i,:));
end

% Code for only a fit to the data and no projections are considered.
Xadj1 = X((21/h+1):1:length(time),:);
Yadj1 = Y((7/h+1):1:(51/h+1),:);
IncidenceAdj1 = Incidence(:,(7/h+1):1:(51/h+1));

% Code for a fit to the data together with projections.
Xadj2 = X((7/h+1):1:length(time),:);
Yadj2 = Y((7/h+1):1:length(time),:);
IncidenceAdj2 = Incidence(:,(7/h+1):1:length(time));

% 3D plot of simulation and datapoints
hold on
% mesh(Xadj1,Yadj1,IncidenceAdj1')  % Alternative to only considering data
mesh(Xadj2,Yadj2,IncidenceAdj2') % Plot for the projection
scatter3(agesvec,yearsvec,Uvec,25,'k','filled')
xlabel('age')
ylabel('time')
zlabel('Incidence')
hold off
```

A.2 Simulations for Cape Town Metropole Data

```
clear all

h = 0.2;  % h is the grid size
begin_year = 1996;
age = 0:h:80;
time = begin_year:h:2015;
[X,Y] = meshgrid(age,time);

% Variable which runs the code for a number of times to provide convergence
numruns = 100;

% Declare all the classes and incidence by means of zero matrices
S = zeros(length(age),length(time));
V = zeros(length(age),length(time));
L = zeros(length(age),length(time));
I = zeros(length(age),length(time));
R = zeros(length(age),length(time));
Incidence = zeros(length(age),length(time));

% Proportion of Susceptibles vaccinated initially
frac = 0.8;


% Boundary Condition
S(1,:) = (2550000/15)*0.29*exp(0.029*(time-1996));

% Initial Conditions
S(:,1) = (1-frac)*183140*(1./(1 + 0.0185*age.^2));
V(:,1) = (frac)*S(:,1);
L(:,1) = 9508*(exp(0.02*age)-1);
I(:,1) = 1.7482*Proportion*(-age + 100).*(age);
R(:,1) = 2.9752*Proportion*(-age + 100).*(age);

%Code to find right population size and proportion of classes
sum(S(:,1) + V(:,1)+ L(:,1)+ I(:,1)+ R(:,1))
[sum(S(:,1)) sum(V(:,1)) sum(L(:,1)) sum(I(:,1)) sum(R(:,1))]
   /sum(S(:,1) + V(:,1)+ L(:,1)+ I(:,1)+ R(:,1))

mu = 0.0003*exp(0.0735*age);

% Ages linked to the corresponding entries of parameter vectors
ages = [0  2  10  20  30  40  50  60  70  80];

% Ages   0  2  10  20  30  40  50  60  70  80

% Parameters vectors corresponding to ages
psi = [0.75 0.675 0.08651251235750 0 0.00000000000000000001
       0 0.00000000000000000001 0 0.00000000000000000001 0];
```

```
p =  [0.8 0.85 0.9 0.88 0.9 0.92 0.93 0.94 0.95 0.97];
r1 = [0.009 0.0075 0.0065 0.9 1.8 2.9 3.25 3.5 3.25 2.9];
r2 = [0.0135 0.00715 0.002715 0.08 0.41 0.415 0.42 0.35 0.32 0.3];
rho = [0.5 0.57 0.65 0.652875 0.65 0.47 0.38 0.39 0.4 0.15];
phi = [0.04 0.06 0.075 0.035 0.2 0.13 0.17 0.21 0.14 0.3];
delta = [0.14 0.07 0.014 0.0315125 0.05125 0.065 0.08 0.1038 0.3 0.4];
sigma = [0.000000005 0.000000015 0.00000005 0.000002 0.0001
         0.00022 0.00025 0.00028 0.0004 0.00045];
theta = [0.05 0.25 0.45 0.55 0.65 0.68 0.72 0.75 0.85 0.95];


% Declare all the transmission matrix by means of a zero matrix
betamatty = zeros(length(time),length(age));
years = [1996 1997 2003:1:2009 2015];

% Specify transmission vectors which entries correspond to:

% Ages      0   2   10   10.4   17    20   30   40   50   60   70   80

beta_1996 = [0.0819702 0.3122 0.102 0.06 0.036407 0.0302 0.020048
             0.0102024 0.0041 0.001402 0.0001 0.00000028051822];
beta_1997 = [0.22541805 0.85855 0.2805 0.165 0.10011925 0.08305 0.055132
             0.0280566 0.011275 0.0038555 0.000275 0.00000077142511];
beta_2003 = [0.50985 2.1230985  1.201539865 0.32031 0.2037  0.235412
             0.1185 0.0832751 0.0397 0.0192 0.0155 0.01237];
beta_2004 = [0.509851 2.045107130  1.27 0.3 0.236 0.24125 0.136275
             0.096 0.04350001 0.0215 0.018675 0.02285];
beta_2005 = [0.50985 2.120530985  1.34 0.31 0.2182034 0.25 0.155
             0.1145 0.0563251 0.47202627 0.030135 0.0401325];
beta_2006 = [0.6509851 2.68413  1.7345 0.3 0.25182035 0.273183 0.1775
             0.139 0.0765 0.03671 0.03675 0.0657];
beta_2007 = [0.50985 2.625 1.6751 0.31 0.236 0.25 0.183215 0.1525
             0.084 0.042151 0.044725 0.103251];

% Ages      0   2   10   10.4   17    20   30   40   50   60   70   80

beta_2008 = [0.7509851 2.870130 1.78 0.3 0.22037 0.23 0.199 0.1775
             0.1 0.05 0.0534862 0.14435];
beta_2009 = [0.750985 2.8013  1.795 0.31 0.182038 0.218 0.20125 0.185725
             0.1125 0.062 0.060885 0.19751];
beta_2015 = [0.563238825 2.6 1.57 0.225 0.13652775 0.195 0.173 0.25
             0.1875 0.12375 0.11427 0.605814763425];


% Assigning the parameter vectors to their respective matrices

bmatty = vertcat(beta_1996,beta_1997,beta_2003,beta_2004,beta_2005,
                 beta_2006,beta_2007,beta_2008,beta_2009,beta_2015);

paramatty = vertcat(psi, p, r1, r2, rho, phi, delta, sigma, theta);
```

```
% Code for interpolating parameters linearly between different ages
size_paramatty = size(paramatty);
nrows_paramatty = size_paramatty(1);
paramdiffy = zeros(nrows_paramatty,length(ages)-1);
param = zeros(nrows_paramatty,length(age));

for i = 1:1:nrows_paramatty
    for j = 1:1:(length(ages)-1)
        paramdiffy(i,j) = ((paramatty(i,j+1) - paramatty(i,j))*h)/
        (ages(j+1) - ages(j));
    end
end

for k = 1:1:(nrows_paramatty)
    vect1 = [];
    for m = 1:1:(length(ages)-2)
        vect1 = horzcat(vect1,paramatty(k,m):paramdiffy(k,m):
        (paramatty(k,m+1)-paramdiffy(k,m)));
    end
    vect1 = horzcat(vect1,paramatty(k,(length(ages)-1)):
    paramdiffy(k,(length(ages)-1)):paramatty(k,length(ages)));
    param(k,:) = vect1;
end

psi = param(1,:);
p = param(2,:);
r1 = param(3,:);
r2 = param(4,:);
rho = param(5,:);
phi = param(6,:);
delta = param(7,:);
sigma = param(8,:);
theta = param(9,:);

ages = [0 2 10 10.4 17 20 30 40 50 60 70 80];
bdiffy = zeros(length(years),length(ages)-1);

for i = 1:1:length(years)
    for j = 1:1:(length(ages)-1)
        bdiffy(i,j) = ((bmatty(i,j+1) - bmatty(i,j))*h)/(ages(j+1) - ages(j));
    end
end

% Variable for converting a year number into a year since the initial year
yearsad = years - begin_year;

% Code for interpolating transmission parameters linearly between different ages and times
for k = 1:1:length(years)
    vect1 = [];
    for m = 1:1:(length(ages)-2)
        vect1 = horzcat(vect1,bmatty(k,m):bdiffy(k,m):
```

```
        (bmatty(k,m+1)-bdiffy(k,m)));
    end
    vect1 = horzcat(vect1,bmatty(k,(length(ages)-1)):
    bdiffy(k,(length(ages)-1)):bmatty(k,length(ages)));
    betamatty(((yearsad(k)/h) + 1),:) = vect1;
end

bmdiffy = zeros(length(years)-1,length(age));

for k = 1:1:length(age)
    for n = 1:1:(length(years)-1)
         bmdiffy(n,k) = ((betamatty(yearsad(n+1)/h + 1,k) -
         betamatty(yearsad(n)/h + 1,k))*h)/(yearsad(n+1)-yearsad(n));
    end
end

for k = 1:1:length(age)
    vect = [];
    for m = 1:1:(length(yearsad)-2)
        vect = horzcat(vect,betamatty(yearsad(m)/h + 1,k):bmdiffy(m,k):
        betamatty(yearsad(m+1)/h + 1,k)-bmdiffy(m,k));
    end
    vect = horzcat(vect,betamatty(yearsad(length(yearsad)-1)/h + 1,k):
    bmdiffy(length(yearsad)-1,k):betamatty(yearsad(length(yearsad))/h + 1,k))';
    betamatty(:,k) = vect';
end

betamatty = betamatty';


ages = [0 2 10 20 30 40 50 60 70 80];

%Double for loop of the Infectives Class'es first guess.
for j = 2:1:length(time)
    for i = 2:1:length(age)
        lambda = betamatty(i-1,j)*I(i-1,j);
        S(i,j) = (S(i,j-1) + S(i-1,j))/(2+h*(lambda + mu(i-1)+psi(i-1)));
        V(i,j) = (h*psi(i-1)*S(i,j) + V(i,j-1) + V(i-1,j))/
        (2+h*(theta(i-1)*lambda + mu(i-1)));
        R(i,j) = (h*rho(i-1)*I(i-1,j) + R(i,j-1) + R(i-1,j))/
        (2+h*(r2(i-1)*lambda + phi(i-1) + mu(i-1)));
        L(i,j) = (h*lambda*(p(i-1)*S(i,j) + theta(i-1)*V(i,j)+ r2(i-1)*R(i,j)) +
        L(i,j-1)+ L(i-1,j))/(2+h*(r1(i-1)*lambda + sigma(i-1) + mu(i-1)));
        I(i,j) = (h*(lambda*(1-p(i-1))*S(i,j) + (lambda*r1(i-1) + sigma(i-1))*
        L(i,j)+ phi(i-1)*R(i,j)) + I(i,j-1)+ I(i-1,j))/
        (2+h*(delta(i-1) + rho(i-1) + mu(i-1)));
    end
end

% Triple for loop of the convergence for the system of difference equations.
for runs = 1:1:numruns
```

```
    for j = 2:1:length(time)
        for i = 2:1:length(age)
        lambda = betamatty(i,j)*I(i,j);
        S(i,j) = (S(i,j-1) + S(i-1,j))/(2+h*(lambda + mu(i)+psi(i)));
        V(i,j) = (h*psi(i)*S(i,j) + V(i,j-1) + V(i-1,j))/
        (2+h*(theta(i)*lambda + mu(i)));
        R(i,j) = (h*rho(i)*I(i,j) + R(i,j-1) + R(i-1,j))/
        (2+h*(r2(i)*lambda + phi(i) + mu(i)));
        L(i,j) = (h*lambda*(p(i)*S(i,j) + theta(i)*V(i,j)+ r2(i)*R(i,j)) +
        L(i,j-1)+ L(i-1,j))/(2+h*(r1(i)*lambda + sigma(i) + mu(i)));
        I(i,j) = (h*(lambda*(1-p(i))*S(i,j) + (lambda*r1(i) + sigma(i))*
        L(i,j)+ phi(i)*R(i,j)) + I(i,j-1)+ I(i-1,j))/
        (2+h*(delta(i) + rho(i) + mu(i)));
        end
    end
end

% Double for loop to fill out the matrix of incidence.
for j = 2:1:length(time)
    for i = 2:1:length(age)
        Incidence(i,j) = ((1-p(i))*betamatty(i,j)*I(i,j)*S(i,j) + sigma(i)*L(i,j) +
        r1(i)*betamatty(i,j)*I(i,j)*L(i,j) + phi(i)*R(i,j))/(S(i,j) + L(i,j) + R(i,j));

    end
end

% Years and ages related to the Cape Town Metropole Data
years = [2003:1:2009];

ages = [2 10 20 30 40 50 60 70 80];
U = zeros(length(years),length(ages));

% Cape Town Metropole Data
U(1,1:9) = [1812 975 3529 5766 4325 2180 859 257 81];
U(2,1:9) = [1791 1004 3771 6367 4637 2183 855 255 100];
U(3,1:9) = [1929 1009 4115 6998 5151 2596 939 344 119];
U(4,1:9) = [2506 1209 4772 7784 5859 3197 1171 357 135];
U(5,1:9) = [2516 1096 4656 7886 6019 3230 1215 373 148];
U(6,1:9) = [2830 1120 4564 8455 6589 3543 1299 387 147];
U(7,1:9) = [2847 1118 4591 8510 6498 3682 1448 386 147];

% Length of age intervals
int_length = [5 10*ones(1,8)];

% Divide each data entry by its corresponding age interval
for i = 1:1:length(years)
      U(i,1:9) =  U(i,1:9)./int_length;
end


% Code for scatterplot
```

```
yearsmat = zeros(length(years),length(ages));
for i = 1:1:length(years)
   yearsmat(i,:) = ones(1,length(ages))*years(i);
end

Uvec = [];
agesvec = [];
yearsvec = [];
for i = 1:1:length(years)
    Uvec = horzcat(Uvec,U(i,:));
    agesvec = horzcat(agesvec,ages);
    yearsvec = horzcat(yearsvec,yearsmat(i,:));
end

% Code for only a fit to the data and no projections are considered.
Xadj1 = X((13/h+1):1:length(time),:);
Yadj1 = Y((7/h+1):1:(13/h+1),:);
IncidenceAdj1 = Incidence(:,(7/h+1):1:(13/h+1));

% Code for a fit to the data together with projections.
Xadj2 = X((7/h+1):1:length(time),:);
Yadj2 = Y((7/h+1):1:length(time),:);
IncidenceAdj2 = Incidence(:,(7/h+1):1:length(time));

figure(1)
hold on
% mesh(Xadj1,Yadj1,IncidenceAdj1')  % Alternative to only considering data
mesh(Xadj2,Yadj2,IncidenceAdj2') % Plot for the projection
scatter3(agesvec,yearsvec,Uvec,25,'k','filled')
xlabel('age')
ylabel('time')
zlabel('Incidence')
hold off
```

A.3 Raw Data: City of Cape Town Health Department

WESTERN CAPE Tuberculosis Programme

Case Finding Report
Report on New and Re-treatment Cases of Tuberculosis
CAPE METRO Quarter 1-4 of 2003

TB Cases	Pulmonary				EP	Total	%
	Smear +	Smear -	No Smear	Total			
New Cases	7536	1700	2440	11676	3003	14679	74.2%
Relapses	2255	1307	165	3727	391	4118	20.8%
After default	603	185	40	828	50	878	4.4%
After failure	69	23	8	100	8	108	0.5%
All other ReRx cases	-	-	-	-	-	-	0.0%
Total	10463	3215	2653 *	16331	3452	19783	100.0%
%	64.1%	19.7%	16.2%	100.0%			
% Pulmonary and EP out of all TB Cases				82.6 %	17.4 %		
% Smear + as proportion of Total PTB : **64.1%**							
* of which children aged 0 - 7: 1962							
This report excludes - patient(s) that died before treatment started							
Rate per 100,000 of population (Population = 0):							
All TB Cases: -							
New Smear Positive PTB Cases: -							

TB Cases		0-4	5-14	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	%
All TB Cases	M	953	452	1631	3248	2821	1502	572	141	47	11367	57.5%
	F	859	523	1898	2518	1504	678	287	116	34	8417	42.5%
	TOTAL	1812	975	3529	5766	4325	2180	859	257	81	19784	100.0%
	%	9.2%	4.9%	17.8%	29.1%	21.9%	11.0%	4.3%	1.3%	0.4%	100.0%	
New Smear Pos Cases	M	16	49	947	1509	1147	586	199	53	15	4521	60.0%
	F	6	107	978	963	569	224	119	39	10	3015	40.0%
	TOTAL	22	156	1925	2472	1716	810	318	92	25	7536	100.0%
	%	0.3%	2.1%	25.5%	32.8%	22.8%	10.7%	4.2%	1.2%	0.3%	100.0%	
Re-Treat Smear Pos Cases	M	10	2	194	645	697	344	119	12	3	2026	69.2%
	F	-	1	171	312	246	116	41	14	-	901	30.8%
	TOTAL	10	3	365	957	943	460	160	26	3	2927	100.0%
	%	0.3%	0.1%	12.5%	32.7%	32.2%	15.7%	5.5%	0.9%	0.1%	100.0%	
- record(s) with missing age												

Bacteriological Coverage = $\frac{\text{Total smear pos + Total smear neg}}{\text{Total PTB minus children 0-7 yrs}}$ = **95.2%**

WESTERN CAPE Tuberculosis Programme

Case Finding Report
Report on New and Re-treatment Cases of Tuberculosis
CAPE METRO Quarter 1-4 of 2004

TB Cases	Pulmonary				EP	Total	%
	Smear +	Smear -	No Smear	Total			
New Cases	8325	1822	2562	12709	2995	15704	74.9%
Relapses	2503	1231	191	3925	407	4332	20.7%
After default	604	130	29	763	64	827	3.9%
After failure	68	20	7	95	5	100	0.5%
All other ReRx cases	-	-	-	-	-	-	0.0%
Total	11500	3203	2789 *	17492	3471	20963	100.0%
%	65.7%	18.3%	15.9%	100.0%			
% Pulmonary and EP out of all TB Cases				83.4 %	16.6 %		
% Smear + as proportion of Total PTB : **65.7%**							
* of which children aged 0 - 7: 2001							
This report excludes - patient(s) that died before treatment started							
Rate per 100,000 of population (Population = 0):							
All TB Cases: -							
New Smear Positive PTB Cases: -							

TB Cases		0-4	5-14	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	%
All TB Cases	M	935	475	1721	3455	3005	1481	561	139	61	11833	56.4%
	F	856	529	2050	2912	1632	702	294	116	39	9130	43.6%
	TOTAL	1791	1004	3771	6367	4637	2183	855	255	100	20963	100.0%
	%	8.5%	4.8%	18.0%	30.4%	22.1%	10.4%	4.1%	1.2%	0.5%	100.0%	
New Smear Pos Cases	M	8	62	1046	1639	1273	603	225	45	20	4921	59.1%
	F	4	121	1085	1177	599	250	113	43	12	3404	40.9%
	TOTAL	12	183	2131	2816	1872	853	338	88	32	8325	100.0%
	%	0.1%	2.2%	25.6%	33.8%	22.5%	10.2%	4.1%	1.1%	0.4%	100.0%	
Re-Treat Smear Pos Cases	M	4	6	213	657	767	362	104	22	6	2141	67.4%
	F	3	7	195	368	295	112	36	13	5	1034	32.6%
	TOTAL	7	13	408	1025	1062	474	140	35	11	3175	100.0%
	%	0.2%	0.4%	12.9%	32.3%	33.4%	14.9%	4.4%	1.1%	0.3%	100.0%	
- record(s) with missing age												

Bacteriological Coverage = $\frac{\text{Total smear pos + Total smear neg}}{\text{Total PTB minus children 0-7 yrs}}$ = **94.9%**

WESTERN CAPE Tuberculosis Programme

Case Finding Report
Report on New and Re-treatment Cases of Tuberculosis
CAPE METRO Quarter 1-4 of 2005

TB Cases	Pulmonary				EP	Total	%
	Smear +	Smear -	No Smear	Total			
New Cases	8842	2348	2845	14035	3012	17047	73.5%
Relapses	2751	1457	203	4411	508	4919	21.2%
After default	799	162	62	1023	71	1094	4.7%
After failure	95	29	9	133	7	140	0.6%
All other ReRx cases	-	-	-	-	-	-	0.0%
Total	12487	3996	3119 *	19602	3598	23200	100.0%
%	63.7%	20.4%	15.9%	100.0%			
% Pulmonary and EP out of all TB Cases				84.5 %	15.5 %		
% Smear + as proportion of Total PTB : **63.7%**							
* of which children aged 0 - 7: 2189							
This report excludes - patient(s) that died before treatment started							
Rate per 100,000 of population (Population = 0):							
All TB Cases: -							
New Smear Positive PTB Cases: -							

TB Cases		0-4	5-14	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	%
All TB Cases	M	1038	506	1895	3776	3301	1773	631	191	61	13172	56.8%
	F	891	503	2220	3222	1850	823	308	153	58	10028	43.2%
	TOTAL	1929	1009	4115	6998	5151	2596	939	344	119	23200	100.0%
	%	8.3%	4.3%	17.7%	30.2%	22.2%	11.2%	4.0%	1.5%	0.5%	100.0%	
New Smear Pos Cases	M	7	69	1117	1729	1269	695	226	75	24	5211	58.9%
	F	2	96	1186	1256	598	297	115	63	18	3631	41.1%
	TOTAL	9	165	2303	2985	1867	992	341	138	42	8842	100.0%
	%	0.1%	1.9%	26.0%	33.8%	21.1%	11.2%	3.9%	1.6%	0.5%	100.0%	
Re-Treat Smear Pos Cases	M	-	2	297	746	850	449	128	30	7	2509	68.8%
	F	-	6	212	378	346	136	38	14	6	1136	31.2%
	TOTAL	-	8	509	1124	1196	585	166	44	13	3645	100.0%
	%	0.0%	0.2%	14.0%	30.8%	32.8%	16.0%	4.6%	1.2%	0.4%	100.0%	
- record(s) with missing age												

Bacteriological Coverage = $\frac{\text{Total smear pos + Total smear neg}}{\text{Total PTB minus children 0-7 yrs}}$ = **94.7%**

WESTERN CAPE Tuberculosis Programme

Case Finding Report
Report on New and Re-treatment Cases of Tuberculosis
CAPE METRO Quarter 1-4 of 2006

TB Cases	Pulmonary				EP	Total	%
	Smear +	Smear -	No Smear	Total			
New Cases	9703	3110	3459	16272	3458	19730	73.1%
Relapses	3180	1691	213	5084	558	5642	20.9%
After default	970	322	80	1372	118	1490	5.5%
After failure	77	32	5	114	14	128	0.5%
All other ReRx cases	-	-	-	-	-	-	0.0%
Total	13930	5155	3757 *	22842	4148	26990	100.0%
%	61.0%	22.6%	16.4%	100.0%			
% Pulmonary and EP out of all TB Cases				84.6 %	15.4 %		
% Smear + as proportion of Total PTB : **61.0%**							
* of which children aged 0 - 7: 2769							
This report excludes - patient(s) that died before treatment started							
Rate per 100,000 of population (Population = 0):							
All TB Cases: -							
New Smear Positive PTB Cases: -							

TB Cases		0-4	5-14	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	%
All TB Cases	M	1267	569	2176	4158	3571	2087	772	224	69	14893	55.2%
	F	1239	640	2596	3626	2288	1110	399	133	66	12097	44.8%
	TOTAL	2506	1209	4772	7784	5859	3197	1171	357	135	26990	100.0%
	%	9.3%	4.5%	17.7%	28.8%	21.7%	11.8%	4.3%	1.3%	0.5%	100.0%	
New Smear Pos Cases	M	5	55	1254	1736	1362	774	273	63	22	5544	57.1%
	F	7	138	1337	1306	801	353	154	42	21	4159	42.9%
	TOTAL	12	193	2591	3042	2163	1127	427	105	43	9703	100.0%
	%	0.1%	2.0%	26.7%	31.4%	22.3%	11.6%	4.4%	1.1%	0.4%	100.0%	
Re-Treat Smear Pos Cases	M	2	4	303	853	890	548	171	40	10	2821	66.7%
	F	1	8	230	479	388	223	56	14	7	1406	33.3%
	TOTAL	3	12	533	1332	1278	771	227	54	17	4227	100.0%
	%	0.1%	0.3%	12.6%	31.5%	30.2%	18.2%	5.4%	1.3%	0.4%	100.0%	
- record(s) with missing age												

Bacteriological Coverage = $\frac{\text{Total smear pos + Total smear neg}}{\text{Total PTB minus children 0-7 yrs}}$ = **95.1%**

WESTERN CAPE Tuberculosis Programme

Case Finding Report
Report on New and Re-treatment Cases of Tuberculosis
CAPE METRO Quarter 1-4 of 2007

TB Cases	Pulmonary				EP	Total	%
	Smear +	Smear -	No Smear	Total			
New Cases	9084	3835	3423	16342	3256	19598	72.2%
Relapses	2915	1711	161	4787	492	5279	19.5%
After default	1254	567	96	1917	177	2094	7.7%
After failure	95	28	10	133	11	144	0.5%
All other ReRx cases	1	4	1	6	18	24	0.1%
Total	13349	6145	3691 *	23185	3954	27139	100.0%
%	57.6%	26.5%	15.9%	100.0%			
% Pulmonary and EP out of all TB Cases				85.4 %	14.6 %		
% Smear + as proportion of Total PTB : **57.6%**							
* of which children aged 0 - 7: 2819							
This report excludes - patient(s) that died before treatment started							
Rate per 100,000 of population (Population = 0):							
All TB Cases: -							
New Smear Positive PTB Cases: -							

TB Cases		0-4	5-14	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	%
All TB Cases	M	1348	486	2072	4106	3656	2162	784	213	77	14904	54.9%
	F	1168	610	2584	3780	2363	1068	431	160	71	12235	45.1%
	TOTAL	2516	1096	4656	7886	6019	3230	1215	373	148	27139	100.0%
	%	9.3%	4.0%	17.2%	29.1%	22.2%	11.9%	4.5%	1.4%	0.5%	100.0%	
New Smear Pos Cases	M	9	63	1133	1576	1262	711	256	80	22	5112	56.3%
	F	4	120	1226	1248	771	355	163	63	22	3972	43.7%
	TOTAL	13	183	2359	2824	2033	1066	419	143	44	9084	100.0%
	%	0.1%	2.0%	26.0%	31.1%	22.4%	11.7%	4.6%	1.6%	0.5%	100.0%	
Re-Treat Smear Pos Cases	M	1	3	296	831	918	548	188	24	8	2817	66.0%
	F	-	5	279	501	386	190	61	20	6	1448	34.0%
	TOTAL	1	8	575	1332	1304	738	249	44	14	4265	100.0%
	%	0.0%	0.2%	13.5%	31.2%	30.6%	17.3%	5.8%	1.0%	0.3%	100.0%	
- record(s) with missing age												

Bacteriological Coverage = $\frac{\text{Total smear pos + Total smear neg}}{\text{Total PTB minus children 0-7 yrs}}$ = **95.7%**

WESTERN CAPE Tuberculosis Programme

Case Finding Report
Report on New and Re-treatment Cases of Tuberculosis
CAPE METRO Quarter 1-4 of 2008

TB Cases	Pulmonary				EP	Total	%
	Smear +	Smear -	No Smear	Total			
New Cases	8972	4945	3919	17836	3367	21203	73.3%
Relapses	2194	1486	200	3880	43	3923	13.6%
After default	724	402	91	1217	9	1226	4.2%
After failure	110	45	21	176	2	178	0.6%
All other ReRx cases	858	937	110	1905	499	2404	8.3%
Total	12858	7815	4341 *	25014	3920	28934	100.0%
%	51.4%	31.2%	17.4%	100.0%			
% Pulmonary and EP out of all TB Cases				86.5 %	13.5 %		
% Smear + as proportion of Total PTB : **51.4%**							
* of which children aged 0 - 7: 3141							
This report excludes - patient(s) that died before treatment started							
Rate per 100,000 of population (Population = 0):							
All TB Cases: -							
New Smear Positive PTB Cases: -							

TB Cases		0-4	5-14	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	%
All TB Cases	M	1417	498	2001	4243	3955	2301	863	209	83	15570	53.8%
	F	1413	622	2563	4212	2634	1242	436	178	64	13364	46.2%
	TOTAL	2830	1120	4564	8455	6589	3543	1299	387	147	28934	100.0%
	%	9.8%	3.9%	15.8%	29.2%	22.8%	12.2%	4.5%	1.3%	0.5%	100.0%	
New Smear Pos Cases	M	4	67	1029	1632	1265	700	267	67	18	5049	56.3%
	F	4	115	1155	1299	767	385	128	55	15	3923	43.7%
	TOTAL	8	182	2184	2931	2032	1085	395	122	33	8972	100.0%
	%	0.1%	2.0%	24.3%	32.7%	22.6%	12.1%	4.4%	1.4%	0.4%	100.0%	
Re-Treat Smear Pos Cases	M	-	-	260	713	814	515	167	27	11	2507	64.5%
	F	-	5	227	477	386	204	53	19	8	1379	35.5%
	TOTAL	-	5	487	1190	1200	719	220	46	19	3886	100.0%
	%	0.0%	0.1%	12.5%	30.6%	30.9%	18.5%	5.7%	1.2%	0.5%	100.0%	
- record(s) with missing age												

Bacteriological Coverage = $\frac{\text{Total smear pos + Total smear neg}}{\text{Total PTB minus children 0-7 yrs}}$ = **94.5%**

WESTERN CAPE Tuberculosis Programme

Case Finding Report
Report on New and Re-treatment Cases of Tuberculosis
CAPE METRO Quarter 1-4 of 2009

TB Cases	Pulmonary				EP	Total	%
	Smear +	Smear -	No Smear	Total			
New Cases	8493	5652	3862	18007	3580	21587	73.9%
Relapses	1658	1166	132	2956	-	2956	10.1%
After default	538	295	59	892	-	892	3.1%
After failure	102	68	9	179	-	179	0.6%
All other ReRx cases	1295	1508	170	2973	640	3613	12.4%
Total	12086	8689	4232 *	25007	4220	29227	100.0%
%	48.3%	34.7%	16.9%	100.0%			
% Pulmonary and EP out of all TB Cases				85.6 %	14.4 %		
% Smear + as proportion of Total PTB : **48.3%**							
* of which children aged 0 - 7: 3119							
This report excludes - patient(s) that died before treatment started							
Rate per 100,000 of population (Population = 0):							
All TB Cases: -							
New Smear Positive PTB Cases: -							

TB Cases		0-4	5-14	15-24	25-34	35-44	45-54	55-64	65-74	>74	Total	%
All TB Cases	M	1454	546	2070	4215	3986	2428	973	219	78	15969	54.6%
	F	1393	572	2521	4295	2512	1254	475	167	69	13258	45.4%
	TOTAL	2847	1118	4591	8510	6498	3682	1448	386	147	29227	100.0%
	%	9.7%	3.8%	15.7%	29.1%	22.2%	12.6%	5.0%	1.3%	0.5%	100.0%	
New Smear Pos Cases	M	5	50	1066	1557	1205	707	278	50	18	4936	58.1%
	F	10	113	1091	1159	653	323	136	48	24	3557	41.9%
	TOTAL	15	163	2157	2716	1858	1030	414	98	42	8493	100.0%
	%	0.2%	1.9%	25.4%	32.0%	21.9%	12.1%	4.9%	1.2%	0.5%	100.0%	
Re-Treat Smear Pos Cases	M	2	3	238	615	766	516	167	30	3	2340	65.1%
	F	1	10	196	424	307	219	77	13	6	1253	34.9%
	TOTAL	3	13	434	1039	1073	735	244	43	9	3593	100.0%
	%	0.1%	0.4%	12.1%	28.9%	29.9%	20.5%	6.8%	1.2%	0.3%	100.0%	
- record(s) with missing age												

Bacteriological Coverage = $\frac{\text{Total smear pos + Total smear neg}}{\text{Total PTB minus children 0-7 yrs}}$ = **94.9%**

Bibliography

[1] E.A. Abayomi, A. Somers, R. Grewal, G. Sissolak, F. Bassa, D. Maartens, P. Jacobs, C. Stefan, and L.W. Ayers, *Impact of the HIV epidemic and Anti-Retroviral Treatment policy on lymphoma incidence and subtypes seen in the Western Cape of South Africa, 2002-2009: Preliminary findings of the Tygerberg Lymphoma Study Group*, Transfusion and Apheresis Science **44** (2011), 161-166.

[2] L.J. Abu-Raddad, L. Sabatelli, J.T. Achterberg, J.D. Sugimoto, I.M. Longini, Jr. C. Dye, and M.E. Halloran, *Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics*, Proceedings of the National Academy of Sciences **Early Edition** (2009), 1-6.

[3] J.P. Aparicio, A.F. Capurro, and C. Castillo-Chavez, *Transmission and Dynamics of Tuberculosis on Generalized Households*, Journal of Theoretical Biology **206** (2000), 327-341.

[4] N. Bacaër, R. Ouifki, C. Pretorius, R. Wood, and B. Williams, *Modeling the joint epidemics of TB and HIV in a South African township*, Journal of Mathematical Biology (2007), 1-37.

[5] N. Bacaër, C. Pretorius and B. Auvert, *An Age-Structured Model for the Potential Impact of Generalized Access to Antiretrovirals on the South African HIV Epidemic*, Bulletin of Mathematical Biology **72** (2010), 2180-2198.

[6] M. Bakir, D.P.S. Dosanjh, J.J. Deeks, A. Soysal, K.A. Millington, S. Efe, Y. Aslan, D. Polat, N. Kodalli, A. Yagci, I. Barlan, N. Bahceciler, E.E. Demiralp, and A. Lalvani, *Use of T Cell-Based Diagnosis of Tuberculosis Infection to Optimize Interpretation of*

Tuberculin Skin Testing in Child Tuberculosis Contacts, Clinical Infectious Diseases **48** (2009), 302-312.

[7] S.M. Blower and H. Dowlatabadi, *Sensitivity and Uncertainty Analysis of Complex Models of Disease Transmission: An HIV Model, as an Example*, International Statistical Review **62** (1994), 229-243.

[8] M.W. Borgdorff, K. Floyd, and J.F. Broekmans, *Interventions to reduce tuberculosis mortality and transmission in low- and middle-income countries*, Bulletin of the World Health Organization **80** (2002), 217-227.

[9] R. Breban and S. Blower, *Letter to the Editor*, Journal of Theoretical

[10] N. Bacaër, C. Pretorius and B. Auvert, *An Age-Structured Model for the Potential Impact of Generalized Access to Antiretrovirals on the South African HIV Epidemic*, Bulletin of Mathematical Biology **72** (2010), 2180-2198. Biology **235** (2005), 151-152.

[11] E. Brooks-Pollock, T. Cohen, and M.Murray, *The Impact of Realistic Age Structure in Simple Models of Tuberculosis Transmission*, PLoS ONE **5** (2010), e8479.

[12] C. Castillo-Chavez and Z. Feng, *Global stability of an age-structure model for TB and its applications to optimal vaccination strategies*, Mathematical Biosciences **151** (1998), 135-154.

[13] City of Cape Town: Strategy and Planning, Strategic Information and GIS, *Discussion Paper: Demographics Scenario*, (2010), http://www.capetown.gov.za/en/Pages/Results.aspx?k=DemographicsDiscussionPaperAugust2010.pdf

[14] T. Cohen, C. Colijn, B. Finklea, and M. Murray, *Exogenous re-infection and the dynamics of tuberculosis epidemics: local effects in a network model of transmission*, The Journal Royal Society Interface **4** (2007), 523-531.

[15] A. Deaton and C. Paxson, *Growth, demographic structure, and national saving in Taiwan*, Research Program in Development Studies, Princeton University (1999), www.princeton.edu/rpds/papers/pdfs/deaton_paxson_growth_taiwan.pdf

[16] R. Diel, A. Nienhaus, C. Lange, K. Meywald-Walter, M. Forßbohm, and T. Schaberg, *Tuberculosis contact investigation with a new, specific blood test in a low-incidence population containing a high proportion of BCG-vaccinated persons*, Respiratory Research **7** (2006), 1-9.

[17] Z. Feng, C. Castillo-Chavez, and A.F. Capurro, *A Model for Tuberculosis with Exogenous Reinfection*, Theoretical Population Biology **57** (2000), 235-247.

[18] M. Gabriela, M. Gomes, A.O. Franco, M.C. Gomes, and G.F. Medley, *The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy*, The Royal Society **271** (2004), 617-623.

[19] M. Gabriela, M. Gomes, P. Rodrigues, F.M. Hilker, N.B. Mantilla-Beniers, M. Muehlen, A.C. Paulo and G.F. Medley, *Implications of partial immunity on the prospects for tuberculosis control by post-exposure interventions*, Journal of Theoretical Biology **248** (2007), 608-617.

[20] M. Gabriela, M. Gomes, L.J. White, and G.F. Medley, *Infection, reinfection, and vaccination under suboptimal immune, protection: epidemiological perspectives*, The Royal Society **228** (2004), 539-549.

[21] A.C. Hesseling, M.F. Cotton, T. Jennings, A. Whitelaw, L.F. Johnson, B. Eley, P. Roux, P. Godfrey-Faussett, and H.S. Schaaf, *High Incidence of Tuberculosis among HIV-Infected Infants: Evidence from a South African Population-Based Study Highlights the Need for Improved Tuberculosis Control Strategies*, Clinical Infectious Diseases **48** (2009), 108-114.

[22] http://www.cdc.gov.tw/lp.asp?ctNode=2604&CtUnit=1499&BaseDSD=7&mp=5.

[23] http://www.indexmundi.com/taiwan/population.html.

[24] B.M.N. Kagina1, B. Abel, M. Bowmaker, T.J. Scriba1, S. Gelderbloem, E. Smit, M. Erasmus, N. Nene, G. Walzl, G. Black, G.D. Hussey, A.C. Hesseling, and W.A. Hanekom, *Delaying BCG vaccination from birth to 10 weeks of age may result in an enhanced memory CD4 T cell response*, Vaccine **27** (2009), 5488-5495.

[25] I. Kleinschmidt, A. Pettifor, N. Morris, C. MacPhail, and H. Rees, *Geographic Distribution of Human Immunodeficiency Virus in South Africa*, The American Society of Tropical Medicine and Hygiene **77** (2007), 1163-1169.

[26] S.D. Lawn, L.G. Bekker, K. Middelkoop, L. Myer, and R. Wood, *Impact of HIV Infection on the Epidemiology of Tuberculosis in a Peri-Urban Community in South Africa: The need for Age-Specific Interventions*, Clinical Infectious Diseases **42** (2006), 1040-1047.

[27] S.D. Lawn, R. Wood, and R.J. Wilkinson, *Changing Concepts of "Latent Tuberculosis Infection" in Patients Living with HIV Infection*, Clinical and Developmental Immunology **2011** (2011), 1-9.

[28] L. Liu, X-Q Zhao, and Y. Zhou, *A Tuberculosis Model with Seasonality*, Bulletin of Mathematical Biology **72** (2010), 931-952.

[29] E.F. Long, N.K. Vaidya and M.L. Brandeau, *Controlling Co-Epidemics: Analysis of HIV and Tuberculosis Infection Dynamics*, Operations Research **56** (2008), 1366-1381.

[30] G. Maartens and N. Beyers, *Tuberculosis in the tropics*, Clinics in Chest Medicine **23** (2002), 341-350.

[31] S. Marino, I.B. Hogue, C.J. Ray and D.E. Kirschner, *A methodology for performing global uncertainty and sensitivity analysis in systems biology*, Journal of Theoretical Biology **254** (2008), 178-196.

[32] M. Monney and J-P. Zellweger, *Active and passive screening for tuberculosis in Vaud Canton, Switzerland*, Swiss Medical Weekly **135** (2005), 469-474.

[33] B.M. Murphy, B.H. Singer, and D.E. Kirschner, *On treatment of tuberculosis in heterogeneous populations*, Journal of Theoretical Biology **223** (2003), 391-404.

[34] M. Murray, *Determinants of cluster distribution in the molecular epidemiology of tuberculosis*, Proceedings of the National Academy of Sciences **99** (2002), 1538-1543.

[35] D.P. Pešut, Z.B. Gledović, A.D. Grgurević, L.M. Nagorni-Obradović, and T.N. Adžić, *Tuberculosis Incidence in Elderly in Serbia: Key Trends in Socioeconomic Transition*, Croatian Medical Journal **49** (2008), 807-812.

[36] C.D. Pretorius, *An investigation into joint HIV and TB epidemics in South Africa*, PhD Thesis, Stellenbosch University.

[37] E. Shim, Z. Feng, M. Martcheva, and C. Castillo-Chavez, *An age-structured epidemic model of rotavirus with vaccination*, Journal of Mathematical Biology **?** (2000), 1-25.

[38] B.H. Singer and D.E. Kirschner, *Influence of backward bifurcation on interpretation of R_0 in a model of epidemic tuberculosis with reinfection*, Mathematical Biosciences and Engineering **1** (2004), 81-93.

[39] D.G. Storla, S. Yimer, and G.A. Bjune, *A systematic review of delay in the diagnosis and treatment of tuberculosis*, BioMed Central **8** (2008), 1-9.

[40] TAIWAN historical demographical data of the whole country, 2011, http://www.populstat.info/Asia/taiwanc.htm

[41] P.W. Uys, P.D. van Helden, and J.W. Hargrove, *Tuberculosis reinfection rate as a proportion of total infection rate correlates with the logarithm of the incidence rate: a mathematical model*, The Journal Royal Society Interface **6** (2009), 11-15.

[42] P.W. Uys, R.M. Warren, and P.D. van Helden, *A Threshold Value for the Time Delay to TB Diagnosis*, PLoS ONE **2** (2007), e757.

[43] P.W. Uys, R. Warren, P.D. van Helden, M. Murray, and T.C. Victor, *Potential of Rapid Diagnosis for Controlling Drug-Susceptible and Drug-Resistant Tuberculosis in Communities Where Mycobacterium tuberculosis, Infections Are Highly Prevalent*, Journal of Clinical Microbiology **47** (2009), 1484-1490.

[44] A. van Rie, N. Beyers, R.P. Gie, M. Kunneke, L. Zietsman, and P.R. Donald, *Childhood tuberculosis in an urban population in South Africa: burden and risk factor*, Archives of Disease in Childhood **80** (1999), 433-437.

[45] E. Vynnycky and P.E.M. Fine, *The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection*, Epidemiology and Infection **119** (1997), 183-201.

[46] P.D. Wang, *Epidemiological trends of childhood tuberculosis in Taiwan, 1998-2005*, International Journal of Tuberculosis Lung Diseases, **12** (2008), 250-254.

[47] H.G. Wiker, T. Mustafa, G.A. Bjune and M. Harboe, *Evidence for waning of latency in a cohort study of tuberculosis*, BioMed Central Infectious Diseases **10** (2010), 1-10.

[48] D.M.A. Winkler, *Modelling the Effects of Re-infection on Disease Dynamics*, HonsBSc Thesis (2009), 1-25.

[49] R. Wood, S. Johnstone-Robertson, P. Uys, J. Hargrove, K. Middelkoop, S.D. Lawn, and L-G Bekker, *Tuberculosis Transmission to Young Children in a South African Community: Modeling Household and Community Infection Risks*, Clinical Infectious Diseases **51** (2010), 401-408.

[50] R. Wood, H. Liang, H. Wu, K. Middelkoop, T. Oni, M.X. Rangaka, R.J. Wilkinson, L-G. Bekker, and S.D. Lawn, *Changing prevalence of tuberculosis infection with increasing age in high-burden townships in South Africa*, International Journal of Tuberculosis Lung Diseases **14** (2010), 406-412.

[51] M-C. Yu, K-J. Bai, J-H. Chang, and C-N. Lee, *Age Transition of Tuberculosis Patients in Taiwan, 1957-2001*, Journal of the Formosan Medical Association **105** (2006), 25-30.